

METHODS OF GENERATING CHIMERIC ADENOVIRUSES AND USES FOR SUCH CHIMERIC ADENOVIRUSES

BACKGROUND OF THE INVENTION

5 The presence of humoral immunity (circulating antibodies) to adenovirus capsid proteins is a barrier to the use of adenovirus vectors for gene therapy. The prototype adenovirus vectors that have been developed for gene therapy are based on subgroup C adenoviruses such as that of serotype 5. The prevalence of neutralizing antibodies against subgroup C adenoviruses is generally high in human populations as
10 a result of frequent exposure to these pathogens. This fact is likely to greatly limit the effectiveness of gene therapy vectors based on serotypes such as Ad5.

 Analysis of the nature of the protective antibodies against adenoviruses has indicated that the most important target is the major capsid protein, hexon [Wolfhart (1988) *J. Virol* 62, 2321; Gall *et al.* (1996) *J. Virol.* 70, 2116]. Several efforts have
15 been made to engineer the hexon so as to evade the anti-hexon antibodies by making chimeric adenoviruses harboring hexons from other serotypes [Roy *et al.* (1998) *J. Virol.* 72, 6875; US Patent No 5,922,315; Gall *et al.* (1998) *J. Virol.* 72, 10260; Youil *et al.* (2002) *Hum. Gene Ther.* 13, 311; Wu *et al.* (2002) *J. Virol.* 76, 12775].
 However, this has been largely unsuccessful when exchanges among distant serotypes
20 are attempted.

 Alternatively, investigators have proposed using adenovirus vectors that rarely cause human infections or using adenoviruses from non-human sources. However, the lack of a practical manner in which to produce large numbers of such
25 vectors has proved to be a hindrance to developing such vectors.

SUMMARY OF THE INVENTION

 The present invention provides a method of modifying adenoviruses having capsids, and particularly, including hexons, from serotypes which are not well
30 adapted for growth in cells useful for adenoviral virion production. The method is useful for production of scalable amounts of adenoviruses. The resulting chimeric adenovirus genomes are useful for a variety of purposes which are described herein.

BEST AVAILABLE COPY

The invention further provides novel, isolated, adenovirus SA18 nucleic acid and amino acid sequences, vectors containing same, cell lines containing such SA18 sequences and/or vectors, and uses thereof.

Other aspects and advantages of the present invention will be readily apparent
5 from the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 provides the map of the genome of the simian adenovirus generated by shotgun cloning as described in the examples below.

10 Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus, termed H5S25H5.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides chimeric adenovirus genomes composed of
15 the left terminal end and right terminal end of an adenovirus which can be cultured in the selected host cell, and the internal regions encoding, at a minimum, the capsid proteins of another adenovirus serotype. This invention is particularly advantageous for generating adenoviruses having serotypes which are difficult to culture in a desired cell type. The invention thus permits generation of chimeric adenoviruses
20 vectors of varying serotypes.

In the embodiments illustrated herein, chimeric adenoviruses have been constructed where most structural proteins, and not merely the hexon or fiber, are derived from an adenovirus of an unrelated serotype, thereby preserving the majority of the protein-protein interactions that are involved in capsid assembly. Most of the
25 early genes such as those encoded by the adenovirus E1 and E4 regions that are responsible for transcription regulation and regulation of the host cell cycle, are retained from a different serotype that is known to result in high titer virus generation in the commonly used cell types, such as HEK 293 which supplies the Ad5 E1 proteins in *trans*.

30 In another embodiment, the invention provides novel nucleic acid and amino acid sequences from Ad SA18, which was originally isolated from vervet monkey [ATCC VR-943]. The present invention further provides novel adenovirus vectors and packaging cell lines to produce those vectors for use in the *in vitro* production of

recombinant proteins or fragments or other reagents. The invention further provides compositions for use in delivering a heterologous molecule for therapeutic or vaccine purposes. Such therapeutic or vaccine compositions contain the adenoviral vectors carrying an inserted heterologous molecule. In addition, novel sequences of the invention are useful in providing the essential helper functions required for production of recombinant adeno-associated viral (AAV) vectors. Thus, the invention provides helper constructs, methods and cell lines which use these sequences in such production methods.

The term "substantial homology" or "substantial similarity," when referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 99% of the aligned sequences.

The term "substantial homology" or "substantial similarity," when referring to amino acids or fragments thereof, indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid (or its complementary strand), there is amino acid sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, or a fragment thereof which is at least 8 amino acids, or more desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

The term "percent sequence identity" or "identical" in the context of nucleic acid sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome (*e.g.*, about 36 kbp), the full-length of an open reading frame of a gene, protein, subunit, or enzyme [see, *e.g.*, the tables providing the adenoviral coding regions], or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, *e.g.*, of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, "percent sequence identity" may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least

about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

Identity is readily determined using such algorithms and computer programs as are defined herein at default settings. Preferably, such identity is over the full
5 length of the protein, enzyme, subunit, or over a fragment of at least about 8 amino acids in length. However, identity may be based upon shorter regions, where suited to the use to which the identical gene product is being put.

As described herein, alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as
10 "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities are also used. There are also a number of algorithms known in the art that can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using Fasta, a program in GCG Version 6.1. Fasta provides alignments and
15 percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1, herein incorporated by reference. Similarly programs are available for performing
20 amino acid alignments. Generally, these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program that provides at least the level of identity or alignment as that provided by the referenced algorithms and programs.

25 As used throughout this specification and the claims, the term "comprise" and its variants including, "comprises", "comprising", among other variants, is inclusive of other components, elements, integers, steps and the like. The term "consists of" or "consisting of" are exclusive of other components, elements, integers, steps and the like.

30 Except where otherwise specified, the term "vector" includes any genetic element known in the art which will deliver a target molecule to a cell, including, naked DNA, a plasmid, phage, transposon, cosmids, episomes, viruses, etc.

By "minigene" is meant the combination of a selected heterologous gene and the other regulatory elements necessary to drive translation, transcription and/or expression of the gene product in a host cell.

As used herein, the term "transcomplement" refers to when a gene (gene product) of one adenovirus serotype supplies an adenovirus serotype lacking this gene (gene product) from another serotype with the missing function. For example, human adenovirus serotype 5 E1a and E1b functions are known to transcomplement E1-deleted chimpanzee adenovirus Pan 9. Similarly, the inventors have found that human Ad5 E1 transcomplements E1-deleted chimpanzee adenovirus serotypes Pan5, Pan6, Pan7, and simian adenovirus serotypes SV1, SV25 and SV39. Other examples of transcomplementing serotypes include human Ad5 and human Ad2, Ad3, Ad4, Ad5, Ad7, and Ad12.

The term "functionally deleted" or "functional deletion" means that a sufficient amount of the gene region is removed or otherwise damaged, *e.g.*, by mutation or modification, so that the gene region is no longer capable of producing functional products of gene expression. If desired, the entire gene region may be removed. Other suitable sites for gene disruption or deletion are discussed elsewhere in the application.

The term "functional" refers to a product (*e.g.*, a protein or peptide) which performs its native function, although not necessarily at the same level as the native product. The term "functional" may also refer to a gene which encodes and from which a desired product can be expressed.

I. Chimeric Adenoviral Vectors

The compositions of this invention include chimeric adenoviral vectors that deliver a heterologous molecule to cells. For delivery of such a heterologous molecule, the vector can be a plasmid or, preferably, a chimeric adenovirus. The chimeric adenoviruses of the invention include adenovirus DNA from at least two source serotypes, a "donating serotype" and a "parental adenovirus" as described in more detail herein, and a minigene.

Because the adenoviral genome contains open reading frames on both strands, in many instances reference is made herein to 5' and 3' ends of the various regions to avoid confusion between specific open reading frames and gene regions. Thus, when

reference is made herein to the “left” and “right” end of the adenoviral genome, this reference is to the ends of the approximately 36 kb adenoviral genome when depicted in schematic form as is conventional in the art [see, *e.g.*, Horwitz, “Adenoviridae and Their Replication”, in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)]. Thus, as used

5 herein, the “left terminal end” of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme left end of the schematic. Typically, the left end refers to be portion of the genome beginning at map unit 0 and extending to the right to include at least the 5’ inverted terminal repeats (ITRs), and excludes the internal regions of the

10 genome encoding the structural genes. As used herein, the “right terminal end” of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme right end of the schematic. Typically, the right end of the adenoviral genome refers to be portion of the genome ending at map unit 36 and extending to the left to include at least the 3’

15 ITRs, and excludes the internal regions of the genome encoding the structural genes.

A. Adenovirus Regulatory Sequences

1. Serotype

The selection of the adenovirus serotype donating its left terminal end and right terminal end can be readily made by one of skill in the art from

20 among serotypes which can readily be cultured in the desired cell line. Among other factors which may be considered in selecting the serotype of the donating serotype is compatibility with the adenovirus serotype which will be supplying the internal regions at the location at which their sequences are hybridized.

Suitable adenoviruses for donating their left and right termini

25 are available from the American Type Culture Collection, Manassas, Virginia, US (ATCC), a variety of academic and commercial sources, or the desired regions of the donating adenoviruses may be synthesized using known techniques with reference to sequences published in the literature or available from databases (*e.g.*, GenBank, etc.). Examples of suitable donating adenoviruses include, without limitation, human

30 adenovirus serotypes 2, 3, 4, 5, 7, and 12, and further including any of the presently identified human types [see, *e.g.*, Horwitz, “Adenoviridae and Their Replication”, in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)] which can be cultured in the desired cell. Similarly adenoviruses known to infect non-human primates (*e.g.*, chimpanzees,

rhesus, macaque, and other simian species) or other non-human mammals and which grow in the desired cell can be employed in the vector constructs of this invention. Such serotypes include, without limitation, chimpanzee adenoviruses Pan 5 [VR-591], Pan6 [VR-592], Pan7 [VR-593], and C68 (Pan9), described in US Patent No. 6,083,716; and simian adenoviruses including, without limitation SV1 [VR-195]; SV25 [SV-201]; SV35; SV15; SV-34; SV-36; SV-37, and baboon adenovirus [VR-275], among others. The sequences of Pan 5 (also termed C5), Pan 6 (also termed C6), Pan 7 (also termed C7), SV1, SV25, and SV39 have been described [WO 03/046124, published 5 June 2003; and in US Patent Application No. 10/739,096, filed December 19, 2003]], which are incorporated by reference. In the following examples, the human 293 cells and adenovirus type 5 (Ad5), Pan9, and Ad40 are used for convenience. However, one of skill in the art will understand that other cell lines and/or comparable regions derived from other adenoviral strains may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

2. Sequences

The minimum sequences which must be supplied by the adenovirus donating its left terminal end and its right terminal end include the 5' cis-elements and the 3' cis-elements necessary for replication and packaging. Typically, the 5' cis-elements necessary for packaging and replication include the 5' inverted terminal repeat (ITR) sequences (which functions as origins of replication) and the native 5' packaging enhancer domains (that contain sequences necessary for packaging linear Ad genomes and enhancer elements for the E1 promoter). The right end of the adenoviral genome includes the 3' cis-elements (including the ITRs) necessary for packaging and encapsidation. Desirably, the adenovirus serotype donating its left and right termini and/or an adenovirus serotype which transcomplements the serotype of the donating adenovirus, further provides the functions of the necessary adenovirus early genes, including E1 (E1a and E1b), E2 (E2a and E2b), and E4 (including at least the ORF6 region). E3 is not essential and may be deleted as desired, *e.g.*, for insertion of a transgene in this region or to provide space for a transgene inserted in another region (typically for packaging it is desirable for the total adenoviral genome to be under 36 kb).

In certain embodiments, the necessary adenovirus early genes are contained in the chimeric construct of the invention. In other embodiment, one or more of the necessary adenovirus early genes can be provided by the packaging host cell or in *trans*.

5 In general, the chimeric adenovirus of the invention contains regulatory sequences from the donating adenovirus serotype, or a transcomplementing serotype, to provide the chimeric adenovirus with compatible regulatory proteins. Optionally, one or more of the necessary adenoviral structural genes is provided by the adenovirus donating its left terminal and its right terminal end.

10 In certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the adenovirus serotype donating its left and right termini. Where all of these genes are
15 derived from the adenovirus serotype donating the 5' and 3' ITRs, a "pseudotyped" virus is formed. In one embodiment, the chimeric adenovirus contains the left end of the adenovirus genome from the donating serotype, from the 5' ITR through the end of the pol gene (or the pTP). In another embodiment, the chimeric adenovirus contains the left end of the donating adenovirus serotype, from the 5' ITR through the
20 penton. In yet another embodiment, the chimeric adenovirus contains the left end of the donating adenovirus serotype, *e.g.*, through the end of pTP, but contains an ITR from an adenovirus serotype heterologous to the donating adenovirus serotype. Still other embodiments will be readily apparent from the present disclosure.

Optionally, one or more of the genes can be hybrids formed
25 from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins (*e.g.*, without limitation, polymerase, terminal protein, IIIa protein). Suitably, these genes express functional proteins which permit packaging of the adenovirus genes into the capsid. Alternatively, one or more of these proteins (whether hybrid or non-hybrid) can be functionally deleted in the
30 chimeric adenovirus. Where desired, any necessary proteins functionally deleted in the chimeric adenovirus can be expressed in *trans* in the packaging cell.

B. Parental Adenovirus Structural Proteins

1. Serotypes

This invention is particularly well adapted for use in generating chimeric adenoviruses in which the capsid proteins are from a parental
5 adenovirus which does not efficiently grow in a desirable host cell. The selection of the parental adenovirus serotype providing the internal regions can be readily made by one of skill in the art based on the information provided herein.

A variety of suitable adenoviruses can serve as a parental adenovirus supplying the regions encoding the structural (*i.e.*, capsid proteins). Many
10 of such adenoviruses can be obtained from the same sources as described above for the donating adenovirus serotypes. Examples of suitable parental adenovirus serotypes includes, without limitation, human adenovirus serotype 40, among others [see, *e.g.*, Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)], and adenoviruses known to infect non-human primates (*e.g.*,
15 chimpanzees, rhesus, macaque, and other simian species) or other non-human mammals, including, without limitation, chimpanzee adenovirus C1, described in US Patent No. 6,083,716, which is incorporated by reference; simian adenoviruses, and baboon adenoviruses, among others. In addition, the parental adenovirus supplying the internal regions may be from a non-naturally occurring adenovirus serotype, such
20 as may be generated using a variety of techniques known to those of skill in the art.

In one embodiment illustrated herein, a chimeric virus that was constructed was that between the chimpanzee adenoviruses Pan-5 and C1 exhibited a higher titer in human 293 cells than the wild-type parental virus. However, the invention is not limited to the use of these chimpanzee adenoviruses, or to the combination of simian-
25 simian, human-human, or simian-human chimeric adenoviruses. For example, it may be desirable to utilize bovine or canine adenoviruses, or other non-human mammalian adenoviruses which do not naturally infect and/or replicate in human cells.

In the following examples, the human adenovirus type 40 (Ad40) and the chimpanzee adenovirus C1, simian Pan 5 and Ad40, and Pan 5 and
30 simian adenovirus SA18, are used. However, one of skill in the art will understand that other adenoviral serotypes may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

2. Sequences

The parental adenovirus provides to the chimeric construct of the invention its internal regions which includes structural proteins necessary for generating a capsid having the desired characteristics of the parental adenovirus.

5 These desired characteristics include, but are not limited to, the ability to infect target cells and delivery a heterologous transgene, the ability to elude neutralizing antibodies directed to another adenovirus serotype (*i.e.*, avoiding clearance due to cross-reactivity), and/or the ability to infect cells in the absence of an immune response to the chimeric adenovirus. The advantages of such characteristics may be most readily apparent in a regimen which involves repeat delivery of adenoviral
10 vectors. The left and right termini of the parent adenovirus, including at least the 5' ITRs, the E1 region, the E4 region and the 3' ITRs are non-functional and, preferably, completely absent. Optionally, all adenovirus regulatory proteins from this parental adenovirus are non-functional and only the structural proteins (or selected structural
15 proteins) are retained.

At a minimum, the parental adenovirus provides the adenoviral late region encoding the hexon protein. Suitably, the parental adenovirus further provides the late regions encoding the penton and the fiber. In certain embodiments, all of the functional adenoviral late regions, including L1 (encoding
20 52/55 Da, IIIa proteins), L2 (encoding penton, VII, V, Mu proteins), L3 (encoding VI, hexon, Endoprotease), L4 (encoding 100 kD, 33 kD, VIII proteins) and L5 (encoding fiber protein) are supplied by the parental adenovirus. Optionally, one or more of these late gene functions, with the exception of those encoding the hexon, penton and fiber proteins, can be functionally deleted. Any necessary structural proteins may be
25 supplied *in trans*.

Thus, in certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the
30 parental adenovirus donating its internal regions. Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins, as described above.

C. The "Minigene"

Typically, an adenoviral vector of the invention is designed to contain a minigene which may be inserted into the site of a partially deleted, fully deleted (absent), or disrupted adenoviral gene. For example, the minigene may be located in the site of such a functional E1 deletion or functional E3 deletion, or another suitable site.

The methods employed for the selection of the transgene, the cloning and construction of the "minigene" and its insertion into the viral vector are within the skill in the art given the teachings provided herein.

1. The transgene

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a polypeptide, protein, or other product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

The composition of the transgene sequence will depend upon the use to which the adenoviral vector will be put. For example, the adenoviral vector may be used as a helper virus in production of recombinant adeno-associated viruses or in production of recombinant adenoviruses deleted of essential adenoviral gene functions which are supplied by the adenoviral vector, or for a variety of production uses. Alternatively, the adenoviral vector may be used for diagnostic purposes.

One type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc. These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other

spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for
5 beta-galactosidase activity. Where the transgene is GFP or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins,
10 peptides, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, si RNAs, small hairpin RNAs, trans-splicing RNAs, catalytic RNAs, and antisense RNAs. One example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

15 The transgene may be used for treatment, *e.g.*, of genetic deficiencies, as a cancer therapeutic or vaccine, for induction of an immune response, and/or for prophylactic vaccine purposes. As used herein, induction of an immune response refers to the ability of a molecule (*e.g.*, a gene product) to induce a T cell and/or a humoral immune response to the molecule. The invention further includes
20 using multiple transgenes, *e.g.*, to correct or ameliorate a condition caused by a multi-subunit protein. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, *e.g.*, for an immunoglobulin, the platelet-derived growth factor, or a dystrophin protein. In order
25 for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when
30 the size of the DNA encoding each of the subunits is small, *e.g.*, the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, *e.g.*, M.L. Donnelly, *et al*, *J. Gen.*

Virol., 78(Pt 1):13-21 (Jan 1997); Furler, S., *et al*, *Gene Ther.*, 8(11):864-873 (June 2001); Klump H., *et al.*, *Gene Ther.*, 8(10):811-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. However, the selected transgene may encode any biologically active product or other product, *e.g.*, a product desirable for study.

Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

2. Vector and Transgene Regulatory Elements

In addition to the major elements identified above for the minigene, the adenoviral vector also includes conventional control elements which are operably linked to the transgene in a manner that permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*i.e.*, Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters which are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, *e.g.*, Boshart *et al*, *Cell*, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 α promoter [Invitrogen].

Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as

temperature, or the presence of a specific physiological state, *e.g.*, acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems

5 have been described and can be readily selected by one of skill in the art. For example, inducible promoters include the zinc-inducible sheep metallothionine (MT) promoter and the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter. Other inducible systems include the T7 polymerase promoter system [WO 98/10088]; the ecdysone insect promoter [No *et al*, *Proc. Natl. Acad.*

10 *Sci. USA*, **93**:3346-3351 (1996)], the tetracycline-repressible system [Gossen *et al*, *Proc. Natl. Acad. Sci. USA*, **89**:5547-5551 (1992)], the tetracycline-inducible system [Gossen *et al*, *Science*, **268**:1766-1769 (1995), see also Harvey *et al*, *Curr. Opin. Chem. Biol.*, **2**:512-518 (1998)]. Other systems include the FK506 dimer, VP16 or p65 using castradiol, diphenol murislerone, the RU486-inducible system [Wang *et al*,

15 *Nat. Biotech.*, **15**:239-243 (1997) and Wang *et al*, *Gene Ther.*, **4**:432-441 (1997)] and the rapamycin-inducible system [Magari *et al*, *J. Clin. Invest.*, **100**:2865-2872 (1997)]. The effectiveness of some inducible promoters increases over time. In such cases one can enhance the effectiveness of such systems by inserting multiple repressors in tandem, *e.g.*, TetR linked to a TetR by an IRES. Alternatively, one can

20 wait at least 3 days before screening for the desired function. One can enhance expression of desired proteins by known means to enhance the effectiveness of this system. For example, using the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE).

In another embodiment, the native promoter for the transgene

25 will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control

30 elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

Another embodiment of the transgene includes a transgene operably linked to a tissue-specific promoter. For instance, if expression in skeletal

muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally occurring promoters (see Li *et al.*, *Nat. Biotech.*, 17:241-245 (1999)).

5 Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake *et al.*, *J. Virol.*, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al.*, *Gene Ther.*, 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot *et al.*, *Hum. Gene Ther.*, 7:1503-14 (1996)), bone osteocalcin (Stein *et al.*, *Mol. Biol. Rep.*, 24:185-96 (1997)); bone sialoprotein (Chen *et al.*, *J. Bone Miner. Res.*, 11:654-64

10 (1996)), lymphocytes (CD2, Hansal *et al.*, *J. Immunol.*, 161:1063-8 (1998); immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen *et al.*, *Cell. Mol. Neurobiol.*, 13:503-15 (1993)), neurofilament light-chain gene (Piccioli *et al.*, *Proc. Natl. Acad. Sci. USA*, 88:5611-5 (1991)), and the neuron-specific vgf gene (Piccioli *et al.*, *Neuron*, 15:373-84 (1995)),

15 among others.

Optionally, vectors carrying transgenes encoding therapeutically useful or immunogenic products may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes

20 (preferably located outside the viral genome to be packaged into a viral particle) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the vector may include an origin of replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, *e.g.*, Sambrook *et al.*, and references cited

25 therein].

These vectors are generated using the techniques and sequences provided herein, in conjunction with techniques known to those of skill in the art. Such techniques include conventional cloning techniques of cDNA such as those described in texts [Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*,

30 Cold Spring Harbor Press, Cold Spring Harbor, NY], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence.

II. Production of the Recombinant Viral Particle

In one embodiment, the invention provides a method of generating recombinant chimeric adenoviral particles in which the capsid of the chimeric adenovirus is of a serotype incapable of efficient growth in the selected host cell. A
5 vector suitable for production of recombinant chimeric adenoviral particles can be generated by direct cloning. Alternatively, such particles can be generated by homologous recombination between a first vector containing the left end of the chimeric adenoviral genome and a second vector containing the right end of the chimeric adenoviral genome. However, any suitable methodology known to those of
10 skill in the art can be readily utilized to generate a vector suitable to generate a production vector, preferably which contains the entire chimeric adenoviral genome, including a minigene. This production vector is then introduced into a host cell in which the adenoviral capsid protein is assembled and the chimeric adenoviral particle assembled as described.

15 The chimeric adenoviruses of the invention include those in which one or more adenoviral genes are absent, or otherwise rendered non-functional. If any of the missing gene functions are essential to the replication and infectivity of the adenoviral particle, these functions are supplied by a complementation (or transcomplementing) cell line or a helper vector expressing these functions during production of the
20 chimeric adenoviral particle.

Examples of chimeric adenoviruses containing such missing adenoviral gene functions include those which are partially or completely deleted in the E1a and/or E1b gene. In such a case, the E1 gene functions can be supplied by the packaging host cell, permitting the chimeric construct to be deleted of E1 gene functions and, if
25 desired, for a transgene to be inserted in this region. Optionally, the E1 gene can be of a serotype which transcomplements the serotype providing the other adenovirus sequences in order to further reduce the possibility of recombination and improve safety. In other embodiments, it is desirable to retain an intact E1a and/or E1b region in the recombinant adenoviruses. Such an intact E1 region may be located in its
30 native location in the adenoviral genome or placed in the site of a deletion in the native adenoviral genome (*e.g.*, in the E3 region).

In another example, all or a portion of the adenovirus delayed early gene E3 may be eliminated from the chimeric adenovirus. The function of adenovirus E3 is

believed to be irrelevant to the function and production of the recombinant virus particle. Chimeric adenovirus vectors may also be constructed having a deletion of at least the ORF6 region of the E4 gene, and more desirably because of the redundancy in the function of this region, the entire E4 region. Still another vector of this invention contains a deletion in the delayed early gene E2a. Similarly, deletions in the intermediate genes IX and IVa₂ may be useful for some purposes. Optionally, deletions may also be made in selected portions of the late genes L1 through L5, as described above.

Other deletions may be made in the other structural or non-structural adenovirus genes. The above-discussed deletions may be used individually, *i.e.*, an adenovirus sequence for use in the present invention may contain deletions in only a single region. Alternatively, deletions of entire genes or portions thereof effective to destroy their biological activity may be used in any combination. For example, in one exemplary vector, the adenovirus sequence may have deletions of the E1 genes and the E4 gene, or of the E1, E2a and E3 genes, or of the E1 and E3 genes, or of E1, E2a and E4 genes, with or without deletion of E3, and so on. As discussed above, such deletions may be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

Examples of suitable transcomplementing serotypes are provided above. The use of transcomplementing serotypes can be particularly advantageous where there is diversity between the Ad sequences in the vector of the invention and the human AdE1 sequences found in currently available packaging cells. In such cases, the use of the current human E1-containing cells prevents the generation of replication-competent adenoviruses during the replication and production process. However, in certain circumstances, it will be desirable to utilize a cell line which expresses the E1 gene products can be utilized for production of an E1-deleted simian adenovirus. Such cell lines have been described. See, *e.g.*, US Patent 6,083,716.

A. Packaging Host Cells

Suitably, the packaging host cell is selected from among cells in which the adenovirus serotype donating the left and right terminal ends of the chimeric genome are capable of efficient growth. The host cells are preferably of mammalian origin, and most preferably are of non-human primate or human origin.

Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549 [ATCC Accession No. CCL 185], 911 cells, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, FJ *et al.*, (1998), *Hum Gene Ther*, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2, HeLa [ATCC Accession No. CCL 2], KB [CCL 17], Detroit [*e.g.*, Detroit 510, CCL 72] and WI-38 [CCL 75] cells, and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. These cell lines are all available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209. Other suitable cell lines may be obtained from other sources. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, *i.e.*, fibroblast, hepatocyte, tumor cell, etc.

As described above, a chimeric adenovirus of the invention can lack one or more functional adenoviral regulatory and/or structural genes which are supplied either by the host cell or in *trans* to effect packaging of the chimeric adenovirus into the viral capsid to generate the viral particle. Thus, the ability of a selected host cell to supply transcomplementing adenoviral sequences may be taken into consideration in selecting a desired host cell.

In one example, the cells are from a stable cell line which expresses adenovirus E1a and E1b functions from a cell line which transcomplements the adenovirus serotype which donates the left and right termini to the chimera of the invention, permitting the chimera to be E1-deleted. Alternatively, where the cell line does not transcomplement the adenovirus donating the termini, E1 functions may be provided by the chimera, or in *trans*.

If desired, one may utilize the sequences provided herein to generate a packaging cell or cell line that expresses, at a minimum, the adenovirus E1 gene from the adenovirus serotype donating the 5' ITR under the transcriptional control of a promoter for expression, or a transcomplementing serotype, in a selected parent cell line. Inducible or constitutive promoters may be employed for this purpose. Examples of such promoters are described in detail elsewhere in this specification. A parent cell is selected for the generation of a novel cell line expressing any desired adenovirus or adenovirus gene, including, *e.g.*, a human Ad5, AdPan5, Pan6, Pan7,

SV1, SV25 or SV39 gene. Without limitation, such a parent cell line may be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], HEK 293, KB [CCL 17], Detroit [*e.g.*, Detroit 510, CCL 72] and WI-38 [CCL 75] cells, among others. Many of these cell lines are all available from the ATCC. Other suitable parent cell lines may be obtained from other sources.

Such E1-expressing cell lines are useful in the generation of chimeric adenovirus E1 deleted vectors. Additionally, or alternatively, the invention provides cell lines that express one or more simian adenoviral gene products, *e.g.*, E1a, E1b, E2a, and/or E4 ORF6, can be constructed using essentially the same procedures for use in the generation of chimeric viral vectors. Such cell lines can be utilized to transcomplement adenovirus vectors deleted in the essential genes that encode those products, or to provide helper functions necessary for packaging of a helper-dependent virus (*e.g.*, adeno-associated virus). The preparation of a host cell according to this invention involves techniques such as assembly of selected DNA sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook et al., cited above, use of overlapping oligonucleotide sequences of the adenovirus genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired nucleotide sequence.

In still another alternative, the essential adenoviral gene products are provided in *trans* by the adenoviral vector and/or helper virus. In such an instance, a suitable host cell can be selected from any biological organism, including prokaryotic (*e.g.*, bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, FJ et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, *i.e.*, fibroblast, hepatocyte, tumor cell, etc.

B. Helper Vectors

Thus, depending upon the adenovirus gene content of the adenoviral vectors and any adenoviral gene functions expressed from the host cell, a helper vector may be necessary to provide sufficient adenovirus gene sequences necessary to produce an infective recombinant viral particle containing the minigene. See, for example, the techniques described for preparation of a "minimal" human Ad vector in International Patent Application WO96/13597, published May 9, 1996, and incorporated herein by reference. Suitably, these helper vectors may be non-replicating genetic elements, a plasmid, or a virus.

Useful helper vectors contain selected adenovirus gene sequences not present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. In one embodiment, the helper virus is replication-defective and contains a variety of adenovirus genes in addition to the sequences described above. Such a helper vector is desirably used in combination with an E1-expressing cell line.

Helper vectors may be formed into poly-cation conjugates as described in Wu *et al*, *J. Biol. Chem.*, **264**:16985-16987 (1989); K. J. Fisher and J. M. Wilson, *Biochem. J.*, **299**:49 (April 1, 1994). A helper vector may optionally contain a second reporter minigene. A number of such reporter genes are known to the art.

The presence of a reporter gene on the helper virus which is different from the transgene on the adenovirus vector allows both the Ad vector and the helper vector to be independently monitored. This second reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

C. Assembly of Viral Particle and Transfection of a Cell Line

Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 μ g to about 100 μ g DNA, and preferably about 10 to about 50 μ g DNA to about 1×10^4 cells to about 1×10^{13} cells, and preferably about 10^5 cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed above, including transfection, and infection. One or more of the adenoviral genes may be stably integrated into the genome of the

host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently.

Furthermore, the promoters for each of the adenoviral genes may be
5 selected independently from a constitutive promoter, an inducible promoter or a native adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (*i.e.*, by the differentiation state or in replicating or quiescent cells) or by exogenously added factors, for example.

Introduction of the molecules (as plasmids or viruses) into the host
10 cell may also be accomplished using techniques known to the skilled artisan and as discussed throughout the specification. In preferred embodiment, standard transfection techniques are used, *e.g.*, CaPO₄ transfection or electroporation.

Assembly of the selected DNA sequences of the adenovirus (as well
as the transgene and other vector elements) into various intermediate plasmids, and
15 the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques. Such techniques include direct cloning as described [G. Gao *et al*, *Gene Ther.* 2003 Oct; **10**(22):1926-1930; US Patent Publication No. 2003-0092161-A, published May 15, 2003; International Patent Application No. PCT/US03/12405]. Other cloning techniques of cDNA such as
20 those described in texts [Sambrook *et al*, cited above], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence can be utilized. Standard transfection and co-transfection techniques are employed, *e.g.*, CaPO₄ precipitation techniques. Other conventional methods employed include homologous
25 recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

For example, following the construction and assembly of the desired minigene-containing viral vector, the vector is transfected *in vitro* in the presence of an optional helper vector into the packaging cell line. The functions expressed from
30 the plasmid, packaging cell line and helper virus, if any, permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the chimeric viral particles. The current method for producing such virus particles is transfection-based. However, the invention is not limited to such

methods. The resulting chimeric adenoviruses are useful in transferring a selected transgene to a selected cell.

III. Use of the Chimeric Adenovirus Vectors

5 The chimeric adenovirus vectors of the invention are useful for gene transfer to a human or veterinary subject (including, non-human primates, non-simian primates, and other mammals) *in vitro*, *ex vivo*, and *in vivo*.

The recombinant adenovirus vectors described herein can be used as expression vectors for the production of the products encoded by the heterologous genes *in vitro*. For example, the recombinant adenoviruses containing a gene inserted into the location of an E1 deletion may be transfected into an E1-expressing cell line as described above. Alternatively, replication-competent adenoviruses may be used in another selected cell line. The transfected cells are then cultured in the conventional manner, allowing the recombinant adenovirus to express the gene product from the promoter. The gene product may then be recovered from the culture medium by known conventional methods of protein isolation and recovery from culture.

A chimeric adenoviral vector of the invention provides an efficient gene transfer vehicle that can deliver a selected transgene to a selected host cell *in vivo* or *ex vivo* even where the organism has neutralizing antibodies to one or more AAV serotypes. In one embodiment, the rAd and the cells are mixed *ex vivo*; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. These compositions are particularly well suited to gene delivery for therapeutic purposes and for immunization, including inducing protective immunity.

More commonly, the chimeric adenoviral vectors of the invention will be utilized for delivery of therapeutic or immunogenic molecules, as described below. It will be readily understood for both applications that the recombinant adenoviral vectors of the invention are particularly well suited for use in regimens involving repeat delivery of recombinant adenoviral vectors. Such regimens typically involve delivery of a series of viral vectors in which the viral capsids are alternated. The viral capsids may be changed for each subsequent administration, or after a pre-selected number of administrations of a particular serotype capsid (*e.g.*, one, two, three, four

or more). Thus, a regimen may involve delivery of a rAd with a first capsid, delivery with a rAd with a second capsid, and delivery with a third capsid. A variety of other regimens which use the Ad capsids of the invention alone, in combination with one another, or in combination with other Ad serotypes will be apparent to those of skill in the art. Optionally, such a regimen may involve administration of rAd with capsids of non-human primate adenoviruses, human adenoviruses, or artificial (*e.g.*, chimeric) serotypes such as are described herein. Each phase of the regimen may involve administration of a series of injections (or other delivery routes) with a single Ad serotype capsid followed by a series with another Ad serotype capsid. Alternatively, the recombinant Ad vectors of the invention may be utilized in regimens involving other non-adenoviral-mediated delivery systems, including other viral systems, non-viral delivery systems, protein, peptides, and other biologically active molecules.

The following sections will focus on exemplary molecules which may be delivered via the adenoviral vectors of the invention.

A. Ad-Mediated Delivery of Therapeutic Molecules

In one embodiment, the Ad vectors described herein are administered to humans according to published methods for gene therapy. A viral vector of the invention bearing the selected transgene may be administered to a patient, preferably suspended in a biologically compatible solution or pharmaceutically acceptable delivery vehicle. A suitable vehicle includes sterile saline. Other aqueous and non-aqueous isotonic sterile injection solutions and aqueous and non-aqueous sterile suspensions known to be pharmaceutically acceptable carriers and well known to those of skill in the art may be employed for this purpose.

The adenoviral vectors are administered in sufficient amounts to transduce the target cells and to provide sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the retina and other intraocular delivery methods, direct delivery to the liver, inhalation, intranasal, intravenous, intramuscular, intratracheal, subcutaneous, intradermal, rectal, oral and other parenteral routes of administration. Routes of administration may be combined, if

desired, or adjusted depending upon the transgene or the condition. The route of administration primarily will depend on the nature of the condition being treated.

Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective adult human or veterinary dosage of the viral vector is generally in the range of from about 100 μ L to about 100 mL of a carrier containing concentrations of from about 1×10^6 to about 1×10^{15} particles, about 1×10^{11} to 1×10^{13} particles, or about 1×10^9 to 1×10^{12} particles. Dosages will range depending upon the size of the animal and the route of administration. For example, a suitable human or veterinary dosage (for about an 80 kg animal) for intramuscular injection is in the range of about 1×10^9 to about 5×10^{12} particles per mL, for a single site. Optionally, multiple sites of administration may be delivered. In another example, a suitable human or veterinary dosage may be in the range of about 1×10^{11} to about 1×10^{15} particles for an oral formulation. One of skill in the art may adjust these doses, depending the route of administration, and the therapeutic or vaccinal application for which the recombinant vector is employed. The levels of expression of the transgene, or for an immunogen, the level of circulating antibody, can be monitored to determine the frequency of dosage administration. Yet other methods for determining the timing of frequency of administration will be readily apparent to one of skill in the art.

An optional method step involves the co-administration to the patient, either concurrently with, or before or after administration of the viral vector, of a suitable amount of a short acting immune modulator. The selected immune modulator is defined herein as an agent capable of inhibiting the formation of neutralizing antibodies directed against the recombinant vector of this invention or capable of inhibiting cytolytic T lymphocyte (CTL) elimination of the vector. The immune modulator may interfere with the interactions between the T helper subsets (T_{H1} or T_{H2}) and B cells to inhibit neutralizing antibody formation. Alternatively, the immune modulator may inhibit the interaction between T_{H1} cells and CTLs to reduce the occurrence of CTL elimination of the vector. A variety of useful immune modulators and dosages for use of same are disclosed, for example, in Yang *et al.*, *J. Virol.*, **70**(9) (Sept 1996); International Patent Application No. WO96/12406, published May 2, 1996; and International Patent Application No. PCT/US96/03035,

all incorporated herein by reference. Typically, such immune modulators would be selected when the transgene is a therapeutic which requires repeat delivery.

1. Therapeutic Transgenes

Useful therapeutic products encoded by the transgene include

- 5 hormones and growth and differentiation factors including, without limitation, insulin, glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF),
10 erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), transforming growth factor α (TGF α), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-II), any one of the transforming growth factor superfamily, including TGF, activins, inhibins, or any of
15 the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregluin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of
20 semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

- Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, *e.g.*, IL-2, IL-
25 4, IL-12 and IL-18), monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors and, interferons, and, stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without
30 limitation, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins, humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include

complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CF2 and CD59.

Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. The invention encompasses receptors for cholesterol regulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, proteins useful in the regulation of lipids, including, *e.g.*, apolipoprotein (apo) A and its isoforms (*e.g.*, ApoAI), apoE and its isoforms including E2, E3 and E4), SRB1, ABC1, and the scavenger receptor. The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as jun, fos, max, mad, serum response factor (SRF), AP-1, AP2, myb, MyoD and myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, *e.g.*, GATA-3, and the forkhead family of winged helix proteins.

Other useful gene products include, carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetate hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, porphobilinogen deaminase, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-coA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylate, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin cDNA sequence. Other useful gene products include those useful for treatment of hemophilia A (*e.g.*, Factor VIII and its variants, including the light chain and heavy chain of the heterodimer, optionally operably linked by a junction), and the B-domain deleted Factor VIII, see US 6,200,560 and 6,221,349], and useful for treatment of hemophilia B (*e.g.*, Factor IX).

Still other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful in certain immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

Reduction and/or modulation of expression of a gene are particularly desirable for treatment of hyperproliferative conditions characterized by hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGRF. In addition to oncogene products as target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that are associated with autoimmunity including cell receptors and cells which produce self-directed antibodies. T-cell mediated autoimmune diseases include rheumatoid arthritis (RA), multiple sclerosis (MS), Sjögren's syndrome, sarcoidosis, insulin dependent diabetes mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis, vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

The chimeric adenoviral vectors of the invention are particularly well suited for therapeutic regimens in which multiple adenoviral-mediated deliveries of transgenes is desired, *e.g.*, in regimens involving redelivery of the same transgene or in combination regimens involving delivery of other transgenes. Such regimens may involve administration of a chimeric adenoviral vector, followed by re-administration with a vector from the same serotype adenovirus. Particularly desirable regimens involve administration of a chimeric adenoviral vector of the invention, in which the serotype of the viral vector delivered in the first administration differs from the serotype of the viral vector utilized in one or more of the subsequent administrations. For example, a therapeutic regimen involves administration of a chimeric vector and repeat administration with one or more adenoviral vectors of the same or different serotypes. In another example, a therapeutic regimen involves administration of an adenoviral vector followed by repeat administration with a chimeric vector of the invention which differs from the serotype of the first delivered adenoviral vector, and optionally further administration with another vector which is the same or, preferably, differs from the serotype of the vector in the prior administration steps. These regimens are not limited to delivery of adenoviral vectors constructed using the chimeric serotypes of the invention. Rather, these regimens can readily utilize chimeric or non-chimeric vectors of other adenoviral serotypes, which may be of artificial, human or non-human primate, or other mammalian sources, in combination with one or more of the chimeric vectors of the invention. Examples of such serotypes are discussed elsewhere in this document. Further, these therapeutic regimens may involve either simultaneous or sequential delivery of chimeric adenoviral vectors of the invention in combination with non-adenoviral vectors, non-viral vectors, and/or a variety of other therapeutically useful compounds or molecules. The present invention is not limited to these therapeutic regimens, a variety of which will be readily apparent to one of skill in the art.

B. Ad-Mediated Delivery of Immunogenic Transgenes

The adenoviruses of the invention may also be employed as immunogenic compositions. As used herein, an immunogenic composition is a composition to which a humoral (*e.g.*, antibody) or cellular (*e.g.*, a cytotoxic T cell) response is mounted to a transgene product delivered by the immunogenic composition following delivery to a mammal, and preferably a primate. The present

invention provides an Ad that can contain in any of its adenovirus sequence deletions a gene encoding a desired immunogen. Chimeric adenoviruses based on simian or other non-human mammalian primate serotypes are likely to be better suited for use as a live recombinant virus vaccine in different animal species compared to an
5 adenovirus of human origin, but is not limited to such a use. The recombinant adenoviruses can be used as prophylactic or therapeutic vaccines against any pathogen for which the antigen(s) crucial for induction of an immune response and able to limit the spread of the pathogen has been identified and for which the cDNA is available.

10 Such vaccinal (or other immunogenic) compositions are formulated in a suitable delivery vehicle, as described above. Generally, doses for the immunogenic compositions are in the range defined above for therapeutic compositions. The levels of immunity of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in
15 the serum, optional booster immunizations may be desired.

Optionally, a vaccinal composition of the invention may be formulated to contain other components, including, *e.g.* adjuvants, stabilizers, pH adjusters, preservatives and the like. Such components are well known to those of skill in the vaccine art. Examples of suitable adjuvants include, without limitation, liposomes,
20 alum, monophosphoryl lipid A, and any biologically active factor, such as cytokine, an interleukin, a chemokine, a ligands, and optimally combinations thereof. Certain of these biologically active factors can be expressed *in vivo*, *e.g.*, via a plasmid or viral vector. For example, such an adjuvant can be administered with a priming DNA vaccine encoding an antigen to enhance the antigen-specific immune response
25 compared with the immune response generated upon priming with a DNA vaccine encoding the antigen only.

The adenoviruses are administered in "an immunogenic amount", that is, an amount of adenovirus that is effective in a route of administration to transfect the desired cells and provide sufficient levels of expression of the selected gene to
30 induce an immune response. Where protective immunity is provided, the recombinant adenoviruses are considered to be vaccine compositions useful in preventing infection and/or recurrent disease.

Alternatively, or in addition, the vectors of the invention may contain a transgene encoding a peptide, polypeptide or protein which induces an immune response to a selected immunogen. The recombinant adenoviruses of this invention are expected to be highly efficacious at inducing cytolytic T cells and antibodies to the inserted heterologous antigenic protein expressed by the vector.

For example, immunogens may be selected from a variety of viral families. Example of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and human enteroviruses such as hepatitis A virus; and the genera apthoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Another viral family includes the calcivirus family, which encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, Ross River virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus, including Rubella virus. The flaviviridae family includes dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinin encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis. In addition, the human coronaviruses include the putative causative agent of sudden acute respiratory syndrome (SARS). Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutinin-esterase) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the rhabdovirus family, which includes the genera vesiculovirus (*e.g.*, Vesicular

Stomatitis Virus), and the general lyssavirus (*e.g.*, rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein. The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus, may be a suitable source of antigens. The paramyxovirus family includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus), parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the family orthomyxovirus and is a suitable source of antigen (*e.g.*, the HA protein, the NI protein). The bunyavirus family includes the genera bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (puremala is a hemahagin fever virus), nairovirus (Nairobi sheep disease) and various unassigned bungaviruses. The arenavirus family provides a source of antigens against LCM and Lassa fever virus. The reovirus family includes the genera reovirus, rotavirus (which causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever), Lebombo (humans), equine encephalosis, blue tongue.

The retrovirus family includes the sub-family oncorivirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLV I and HTLV II, lentivirinal (which includes human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus, and spumavirinal). Among the lentiviruses, many suitable antigens have been described and can readily be selected. Examples of suitable HIV and SIV antigens include, without limitation the gag, pol, Vif, Vpx, VPR, Env, Tat, Nef, and Rev proteins, as well as various fragments thereof. For example, suitable fragments of the Env protein may include any of its subunits such as the gp120, gp160, gp41, or smaller fragments thereof, *e.g.*, of at least about 8 amino acids in length. Similarly, fragments of the tat protein may be selected. [See, US Patent 5,891,994 and US Patent 6,193,981.] See, also, the HIV and SIV proteins described in D.H. Barouch *et al*, *J. Virol.*, 75(5):2462-2467 (March 2001), and R.R. Amara, *et al*, *Science*, 292:69-74 (6 April 2001). In another example, the HIV and/or SIV immunogenic proteins or peptides may be used to form fusion proteins or other immunogenic molecules. See, *e.g.*, the HIV-1 Tat and/or Nef fusion proteins and

immunization regimens described in WO 01/54719, published August 2, 2001, and WO 99/16884, published April 8, 1999. The invention is not limited to the HIV and/or SIV immunogenic proteins or peptides described herein. In addition, a variety of modifications to these proteins have been described or could readily be made by one of skill in the art. See, *e.g.*, the modified gag protein that is described in US Patent 5,972,596. Further, any desired HIV and/or SIV immunogens may be delivered alone or in combination. Such combinations may include expression from a single vector or from multiple vectors. Optionally, another combination may involve delivery of one or more expressed immunogens with delivery of one or more of the immunogens in protein form. Such combinations are discussed in more detail below.

The papovavirus family includes the sub-family polyomaviruses (BKU and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis. The parvovirus includes family feline parvovirus (feline enteritis), feline panleucopeniavirus, canine parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family alpha herpesvirinae, which encompasses the genera simplexvirus (HSV I, HSV II), varicellovirus (pseudorabies, varicella zoster) and the sub-family beta herpesvirinae, which includes the genera cytomegalovirus (HCMV, muromegalovirus) and the sub-family gamma herpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitt's lymphoma), infectious rhinotracheitis, Marek's disease virus, and rhadinovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus family includes equine arteritis virus and various Encephalitis viruses.

The viruses of the present invention may also carry immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell.

Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci; and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; pseudomonas, acinetobacteria and eikenella; melioidosis; 5 salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (which causes chancroid); brucella; *Francisella tularensis* (which causes tularemia); yersinia (pasteurella); streptobacillus moniliformis and spirillum; Gram-positive bacilli include listeria monocytogenes; erysipelotheirus rhusiopathiae; *Corynebacterium diphtheria* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and 10 bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; treponematoses: yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include actinomycosis; nocardiosis; cryptococcosis, blastomycosis, 15 histoplasmosis and coccidioidomycosis; candidiasis; aspergillosis, and mucormycosis; sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever, Rocky Mountain spotted fever, Q fever, and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; 20 lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections. Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis; toxoplasmosis; *Pneumocystis carinii*; Trichans; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes 25 or flukes; and cestode (tapeworm) infections.

Many of these organisms and/or toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax), 30 *Clostridium botulinum* and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fevers [filoviruses (e.g., Ebola, Marburg), and arenaviruses [e.g., Lassa, Machupo]], all of which are currently classified as Category A agents; *Coxiella burnetii* (Q fever);

Brucella species (brucellosis), *Burkholderia mallei* (glanders), *Burkholderia pseudomallei* (meloidosis), *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), Staphylococcus species and their toxins (enterotoxin B), *Chlamydia psittaci* (psittacosis), water safety threats (e.g., *Vibrio cholerae*, *Cryptosporidium parvum*), Typhus fever (*Richettsia powazekii*), and viral encephalitis (alphaviruses, e.g., Venezuelan equine encephalitis; eastern equine encephalitis; western equine encephalitis); all of which are currently classified as Category B agents; and Nipah virus and hantaviruses, which are currently classified as Category C agents. In addition, other organisms, which are so classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described herein are useful to deliver antigens from these organisms, viruses, their toxins or other by-products, which will prevent and/or treat infection or other adverse reactions with these biological agents.

Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In RA, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and V α -17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In MS, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-7 and V α -10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in MS. In scleroderma, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and V α -16, V α -3C, V α -7, V α -14, V α -15, V α -16, V α -28 and V α -12. Thus, delivery of a chimeric adenovirus that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

C. Ad-Mediated Delivery Methods

The therapeutic levels, or levels of immunity, of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of CD8+ T cell response, or optionally, antibody titers, in the serum, optional booster

immunizations may be desired. Optionally, the adenoviral vectors of the invention may be delivered in a single administration or in various combination regimens, *e.g.*, in combination with a regimen or course of treatment involving other active ingredients or in a prime-boost regimen. A variety of such regimens have been
5 described in the art and may be readily selected.

For example, prime-boost regimens may involve the administration of a DNA (*e.g.*, plasmid) based vector to prime the immune system to a second or further, booster, administration with a traditional antigen, such as a protein or a recombinant virus carrying the sequences encoding such an antigen. See, *e.g.*, WO
10 00/11140, published March 2, 2000, incorporated by reference. Alternatively, an immunization regimen may involve the administration of a chimeric adenoviral vector of the invention to boost the immune response to a vector (either viral or DNA-based) carrying an antigen, or a protein. In still another alternative, an immunization regimen involves administration of a protein followed by booster with a vector
15 encoding the antigen.

In one embodiment, the invention provides a method of priming and boosting an immune response to a selected antigen by delivering a plasmid DNA vector carrying said antigen, followed by boosting with an adenoviral vector of the invention. In one embodiment, the prime-boost regimen involves the expression of
20 multiproteins from the prime and/or the boost vehicle. See, *e.g.*, R.R. Amara, *Science*, 292:69-74 (6 April 2001) which describes a multiprotein regimen for expression of protein subunits useful for generating an immune response against HIV and SIV. For example, a DNA prime may deliver the Gag, Pol, Vif, VPX and Vpr and Env, Tat, and Rev from a single transcript. Alternatively, the SIV Gag, Pol and
25 HIV-1 Env is delivered in a recombinant adenovirus construct of the invention. Still other regimens are described in WO 99/16884 and WO 01/54719.

However, the prime-boost regimens are not limited to immunization for HIV or to delivery of these antigens. For example, priming may involve delivering with a first vector of the invention followed by boosting with a second
30 vector, or with a composition containing the antigen itself in protein form. In one example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or other organism from which the antigen is derived. In another desired embodiment, the prime-boost regimen provides a therapeutic effect that can

be measured using convention assays for detection of the presence of the condition for which therapy is being administered.

The priming composition may be administered at various sites in the body in a dose dependent manner, which depends on the antigen to which the desired immune response is being targeted. The invention is not limited to the amount or situs of injection(s) or to the pharmaceutical carrier. Rather, the regimen may involve a priming and/or boosting step, each of which may include a single dose or dosage that is administered hourly, daily, weekly or monthly, or yearly. As an example, the mammals may receive one or two doses containing between about 10 μg to about 50 μg of plasmid in carrier. A desirable amount of a DNA composition ranges between about 1 μg to about 10,000 μg of the DNA vector. Dosages may vary from about 1 μg to 1000 μg DNA per kg of subject body weight. The amount or site of delivery is desirably selected based upon the identity and condition of the mammal. The dosage unit of the vector suitable for delivery of the antigen to the mammal is described herein. The vector is prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline; isotonic salts solution or other formulations that will be apparent to those skilled in such administration. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The compositions of the invention may be administered to a mammal according to the routes described above, in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes. Optionally, the priming step of this invention also includes administering with the priming composition, a suitable amount of an adjuvant, such as are defined herein.

Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming composition to the mammalian subject. The administration of the boosting composition is accomplished using an effective amount of a boosting composition containing or capable of delivering the same antigen as administered by the priming DNA vaccine. The boosting composition may be composed of a recombinant viral vector derived from the same viral source (*e.g.*, adenoviral sequences of the invention) or from another source. Alternatively, the "boosting composition" can be a composition containing the same antigen as encoded in the priming DNA vaccine, but in the form of a protein or peptide, which

composition induces an immune response in the host. In another embodiment, the boosting composition contains a DNA sequence encoding the antigen under the control of a regulatory sequence directing its expression in a mammalian cell, *e.g.*, vectors such as well-known bacterial or viral vectors. The primary requirements of the boosting composition are that the antigen of the composition is the same antigen, or a cross-reactive antigen, as that encoded by the priming composition.

In another embodiment, the adenoviral vectors of the invention are also well suited for use in a variety of other immunization and therapeutic regimens. Such regimens may involve delivery of adenoviral vectors of the invention simultaneously or sequentially with Ad vectors of different serotype capsids, regimens in which adenoviral vectors of the invention are delivered simultaneously or sequentially with non-Ad vectors, regimens in which the adenoviral vectors of the invention are delivered simultaneously or sequentially with proteins, peptides, and/or other biologically useful therapeutic or immunogenic compounds. Such uses will be readily apparent to one of skill in the art.

IV. Simian Adenovirus 18 Sequences

The invention provides nucleic acid sequences and amino acid sequences of Ad SA18, which are isolated from the other viral material with which they are associated in nature. These sequences are useful in preparing heterologous molecules containing the nucleic acid sequences and amino acid sequences, and regions or fragments thereof as are described herein, viral vectors which are useful for a variety of purposes, including the constructs and compositions, and such methods as are described herein for the chimeric adenoviruses, including, *e.g.*, in host cells for production of viruses requiring adenoviral helper functions, as delivery vehicles for heterologous molecules such as those described herein. These sequences are also useful in generating the chimeric adenoviruses of the invention.

A. Nucleic Acid Sequences

The SA18 nucleic acid sequences of the invention include nucleotides SEQ ID NO: 12, nt 1 to 31967. See, Sequence Listing, which is incorporated by reference herein. The nucleic acid sequences of the invention further encompass the strand which is complementary to the sequences of SEQ ID NO: 12, as well as the RNA and cDNA sequences corresponding to the sequences of these sequences figures

and their complementary strands. Further included in this invention are nucleic acid sequences which are greater than 95 to 98%, and more preferably about 99 to 99.9% homologous or identical to the Sequence Listing. Also included in the nucleic acid sequences of the invention are natural variants and engineered modifications of the sequences provided in SEQ ID NO: 12 and their complementary strands. Such modifications include, for example, labels that are known in the art, methylation, and substitution of one or more of the naturally occurring nucleotides with a degenerate nucleotide.

The invention further encompasses fragments of the sequences of SA18, their complementary strand, cDNA and RNA complementary thereto. Suitable fragments are at least 15 nucleotides in length, and encompass functional fragments, *i.e.*, fragments which are of biological interest. For example, a functional fragment can express a desired adenoviral product or may be useful in production of recombinant viral vectors. Such fragments include the gene sequences and fragments listed in the tables below.

The following tables provide the transcript regions and open reading frames in the simian adenovirus sequences of the invention. For certain genes, the transcripts and open reading frames (ORFs) are located on the strand complementary to that presented in SEQ ID NO: 12. See, *e.g.*, E2b, E4 and E2a. The calculated molecular weights of the encoded proteins are also shown.

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
		start	End	M.W.
ITR		1	180	
E1a	13S	916	1765	27264
	12S	916	1765	24081
E1b	Small T	1874	2380	19423
	Large T	2179	3609	52741
	IX	3678	4079	13701
E2b	IVa2	5478	4126	51295
	Polymerase	13745	5229	128392
	PTP	13745	8597	75358
	Agnoprotein	8007	8705	23610
L1	52/55 kD	10788	11945	43416
	IIIa	11966	13699	63999
L2	Penton	13796	15322	57166
	VII	15328	15873	20352
	V	15920	17050	42020
L3	VI	17348	18154	29222
	Hexon	18257	21010	102912
	Endoprotease	21029	21640	23015

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
2a	DBP	23147	21711	53626
L4	100kD	23175	25541	87538
	22 kD homolog	25204	25797	22206
	33 kD homolog	25204	26025	24263
	VIII	26107	26817	25490
E3	Orf #1	26817	27125	11814
L5	Fiber	27192	29015	65455
E4	Orf 6/7	30169	29067	13768
	Orf 6	30169	29303	33832
	Orf 4	30464	30099	14154
	Orf 3	30816	30466	13493
	Orf 2	31205	30813	14698
	Orf 1	31608	31231	14054
ITR		31788	31967	

The SA18 adenoviral nucleic acid sequences are useful as therapeutic and immunogenic agents and in construction of a variety of vector systems and host cells.

5 Such vectors are useful for any of the purposes described above for the chimeric adenovirus. Additionally, these SA18 sequences and products may be used alone or in combination with other adenoviral sequences or fragments, or in combination with elements from other adenoviral or non-adenoviral sequences. The adenoviral sequences of the invention are also useful as antisense delivery vectors, gene therapy
10 vectors, or vaccine vectors, and in methods of using same. Thus, the invention further provides nucleic acid molecules, gene delivery vectors, and host cells which contain the Ad sequences of the invention.

For example, the invention encompasses a nucleic acid molecule containing simian Ad ITR sequences of the invention. In another example, the invention
15 provides a nucleic acid molecule containing simian Ad sequences of the invention

encoding a desired Ad gene product. Still other nucleic acid molecule constructed using the sequences of the invention will be readily apparent to one of skill in the art, in view of the information provided herein.

5 In one embodiment, the simian Ad gene regions identified herein may be used in a variety of vectors for delivery of a heterologous molecule to a cell. Examples of such molecules and methods of delivery are provided in Section III herein. For example, vectors are generated for expression of an adenoviral capsid protein (or fragment thereof) for purposes of generating a viral vector in a packaging host cell. Such vectors may be designed for expression in trans. Alternatively, such vectors are
10 designed to provide cells which stably contain sequences which express desired adenoviral functions, *e.g.*, one or more of E1a, E1b, the terminal repeat sequences, E2a, E2b, E4, E4ORF6 region.

In addition, the adenoviral gene sequences and fragments thereof are useful for providing the helper functions necessary for production of helper-dependent
15 viruses (*e.g.*, adenoviral vectors deleted of essential functions or adeno-associated viruses (AAV)). For such production methods, the simian adenoviral sequences of the invention are utilized in such a method in a manner similar to those described for the human Ad. However, due to the differences in sequences between the simian adenoviral sequences of the invention and those of human Ad, the use of the
20 sequences of the invention essentially eliminate the possibility of homologous recombination with helper functions in a host cell carrying human Ad E1 functions, *e.g.*, 293 cells, which may produce infectious adenoviral contaminants during rAAV production.

Methods of producing rAAV using adenoviral helper functions have been
25 described at length in the literature with human adenoviral serotypes. See, *e.g.*, US Patent 6,258,595 and the references cited therein. See, also, US Patent 5,871,982; WO 99/14354; WO 99/15685; WO 99/47691. These methods may also be used in production of non-human serotype AAV, including non-human primate AAV serotypes. The simian adenoviral gene sequences of the invention which provide the
30 necessary helper functions (*e.g.*, E1a, E1b, E2a and/or E4 ORF6) can be particularly useful in providing the necessary adenoviral function while minimizing or eliminating the possibility of recombination with any other adenoviruses present in the rAAV-packaging cell which are typically of human origin. Thus, selected genes or open

reading frames of the adenoviral sequences of the invention may be utilized in these rAAV production methods.

Alternatively, recombinant adenoviral simian vectors of the invention may be utilized in these methods. Such recombinant adenoviral simian vectors may include, *e.g.*, a hybrid simian Ad/AAV in which simian Ad sequences flank a rAAV expression cassette composed of, *e.g.*, AAV 3' and/or 5' ITRs and a transgene under the control of regulatory sequences which control its expression. One of skill in the art will recognize that still other simian adenoviral vectors and/or gene sequences of the invention will be useful for production of rAAV and other viruses dependent upon adenoviral helper.

In still another embodiment, nucleic acid molecules are designed for delivery and expression of selected adenoviral gene products in a host cell to achieve a desired physiologic effect. For example, a nucleic acid molecule containing sequences encoding an adenovirus E1a protein of the invention may be delivered to a subject for use as a cancer therapeutic. Optionally, such a molecule is formulated in a lipid-based carrier and preferentially targets cancer cells. Such a formulation may be combined with other cancer therapeutics (*e.g.*, cisplatin, taxol, or the like). Still other uses for the adenoviral sequences provided herein will be readily apparent to one of skill in the art.

In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for *in vitro*, *ex vivo* or *in vivo* delivery of therapeutic and immunogenic molecules, including any of those identified as being deliverable via the chimeric adenoviruses of the invention. For example, the simian Ad genome of the invention can be utilized in a variety of rAd and non-rAd vector systems. Such vectors systems may include, *e.g.*, plasmids, lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adeno-associated viral systems, among others. Selection of these vector systems is not a limitation of the present invention.

The invention further provides molecules useful for production of the simian and simian-derived proteins of the invention. Such molecules which carry polynucleotides including the simian Ad DNA sequences of the invention can be in the form of a vector.

B. Simian Adenoviral Proteins of the Invention

The invention further provides gene products of the above adenoviruses, such as proteins, enzymes, and fragments thereof, which are encoded by the adenoviral nucleic acids of the invention. The invention further encompasses
5 SA18 proteins, enzymes, and fragments thereof, having the amino acid sequences encoded by these nucleic acid sequences which are generated by other methods. Such proteins include those encoded by the open reading frames identified in the tables above, and fragments thereof.

Thus, in one aspect, the invention provides unique simian adenoviral
10 proteins which are substantially pure, *i.e.*, are free of other viral and proteinaceous proteins. Preferably, these proteins are at least 10% homogeneous, more preferably 60% homogeneous, and most preferably 95% homogeneous.

In one embodiment, the invention provides unique simian-derived capsid proteins. As used herein, a simian-derived capsid protein includes any
15 adenoviral capsid protein that contains a SA18 capsid protein or a fragment thereof, as defined above, including, without limitation, chimeric capsid proteins, fusion proteins, artificial capsid proteins, synthetic capsid proteins, and recombinantly capsid proteins, without limitation to means of generating these proteins.

Suitably, these simian-derived capsid proteins contain one or more
20 SA18 regions or fragments thereof (*e.g.*, a hexon, penton, fiber or fragment thereof) in combination with capsid regions or fragments thereof of different adenoviral serotypes, or modified simian capsid proteins or fragments, as described herein. A "modification of a capsid protein associated with altered tropism" as used herein includes an altered capsid protein, *i.e.*, a penton, hexon or fiber protein region, or
25 fragment thereof, such as the knob domain of the fiber region, or a polynucleotide encoding same, such that specificity is altered. The simian-derived capsid may be constructed with one or more of the simian Ad of the invention or another Ad serotypes which may be of human or non-human origin. Such Ad may be obtained from a variety of sources including the ATCC, commercial and academic sources, or
30 the sequences of the Ad may be obtained from GenBank or other suitable sources.

The amino acid sequences of the simian adenoviruses penton proteins of the invention are provided herein. The AdSA18 penton protein is provided in SEQ ID NO: 13. Suitably, any of these penton proteins, or unique fragments thereof, may

be utilized for a variety of purposes. Examples of suitable fragments include the penton having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. Further, the penton protein may be modified for a variety of purposes known to those of skill in the art.

The invention further provides the amino acid sequences of the hexon protein of SA18, SEQ ID NO:14. Suitably, this hexon protein, or unique fragments thereof, may be utilized for a variety of purposes. Examples of suitable fragments include the hexon having N-terminal and/or C-terminal truncations of about 50, 100, 150, 200, 300, 400, or 500 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 14. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. For example, one suitable fragment the loop region (domain) of the hexon protein, designated DE1 and FG1, or a hypervariable region thereof. Such fragments include the regions spanning amino acid residues about 125 to 443; about 138 to 441, or smaller fragments, such as those spanning about residue 138 to residue 163; about 170 to about 176; about 195 to about 203; about 233 to about 246; about 253 to about 264; about 287 to about 297; about 404 to about 430, about 430 to 550, about 545 to 650; of the simian hexon proteins, with reference to SEQ ID NO: 14. Other suitable fragments may be readily identified by one of skill in the art. Further, the hexon protein may be modified for a variety of purposes known to those of skill in the art. Because the hexon protein is the determinant for serotype of an adenovirus, such artificial hexon proteins would result in adenoviruses having artificial serotypes. Other artificial capsid proteins can also be constructed using the chimp Ad penton sequences and/or fiber sequences of the invention and/or fragments thereof.

In one example, it may be desirable to generate an adenovirus having an altered hexon protein utilizing the sequences of a hexon protein of the invention. One suitable method for altering hexon proteins is described in US Patent 5,922,315, which is incorporated by reference. In this method, at least one loop region of the adenovirus hexon is changed with at least one loop region of another adenovirus serotype. Thus, at least one loop region of such an altered adenovirus hexon protein is a simian Ad hexon loop region of the invention. In one embodiment, a loop region

of the SA18 hexon protein is replaced by a loop region from another adenovirus serotype. In another embodiment, the loop region of the SA18 hexon is used to replace a loop region from another adenovirus serotype. Suitable adenovirus serotypes may be readily selected from among human and non-human serotypes, as described herein. SA18 is selected for purposes of illustration only; the other simian Ad hexon proteins of the invention may be similarly altered, or used to alter another Ad hexon. The selection of a suitable serotype is not a limitation of the present invention. Still other uses for the hexon protein sequences of the invention will be readily apparent to those of skill in the art.

The invention further encompasses the fiber proteins of the simian adenoviruses of the invention. The fiber protein of AdSA18 has the amino acid sequence of SEQ ID NO: 15. Suitably, this fiber protein, or unique fragments thereof, may be utilized for a variety of purposes. One suitable fragment is the fiber knob, which spans about amino acids 247 to 425 of SEQ ID NO: 15. Examples of other suitable fragments include the fiber having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 15. Still other suitable fragments include internal fragments. Further, the fiber protein may be modified using a variety of techniques known to those of skill in the art.

The invention further encompasses unique fragments of the proteins of the invention which are at least 8 amino acids in length. However, fragments of other desired lengths can be readily utilized. In addition, the invention encompasses such modifications as may be introduced to enhance yield and/or expression of an SA18 gene product, *e.g.*, construction of a fusion molecule in which all or a fragment of the SA18 gene product is fused (either directly or via a linker) with a fusion partner to enhance. Other suitable modifications include, without limitation, truncation of a coding region (*e.g.*, a protein or enzyme) to eliminate a pre- or pro-protein ordinarily cleaved and to provide the mature protein or enzyme and/or mutation of a coding region to provide a secretable gene product. Still other modifications will be readily apparent to one of skill in the art. The invention further encompasses proteins having at least about 95% to 99% identity to the SA18 proteins provided herein.

As described herein, vectors of the invention containing the adenoviral capsid proteins of the invention are particularly well suited for use in applications in which the neutralizing antibodies diminish the effectiveness of other Ad serotype based vectors, as well as other viral vectors. The rAd vectors of the invention are particularly advantageous in readministration for repeat gene therapy or for boosting immune response (vaccine titers). Examples of such regimens are provided herein.

Under certain circumstances, it may be desirable to use one or more of the SA18 gene products (*e.g.*, a capsid protein or a fragment thereof) to generate an antibody. The term "an antibody," as used herein, refers to an immunoglobulin molecule which is able to specifically bind to an epitope. Thus, the antibodies of the invention bind, preferably specifically and without cross-reactivity, to a SA18 epitope. The antibodies in the present invention exist in a variety of forms including, for example, high affinity polyclonal antibodies, monoclonal antibodies, synthetic antibodies, chimeric antibodies, recombinant antibodies and humanized antibodies. Such antibodies originate from immunoglobulin classes IgG, IgM, IgA, IgD and IgE.

Such antibodies may be generated using any of a number of methods known in the art. Suitable antibodies may be generated by well-known conventional techniques, *e.g.* Kohler and Milstein and the many known modifications thereof. Similarly desirable high titer antibodies are generated by applying known recombinant techniques to the monoclonal or polyclonal antibodies developed to these antigens [see, *e.g.*, PCT Patent Application No. PCT/GB85/00392; British Patent Application Publication No. GB2188638A; Amit *et al.*, 1986 *Science*, 233:747-753; Queen *et al.*, 1989 *Proc. Nat'l. Acad. Sci. USA*, 86:10029-10033; PCT Patent Application No. PCT/WO9007861; and Riechmann *et al.*, *Nature*, 332:323-327 (1988); Huse *et al.*, 1988a *Science*, 246:1275-1281]. Alternatively, antibodies can be produced by manipulating the complementarity determining regions of animal or human antibodies to the antigen of this invention. See, *e.g.*, E. Mark and Padlin, "Humanization of Monoclonal Antibodies", Chapter 4, The Handbook of Experimental Pharmacology, Vol. 113, The Pharmacology of Monoclonal Antibodies, Springer-Verlag (June, 1994); Harlow *et al.*, 1999, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY; Harlow *et al.*, 1989, Antibodies: A Laboratory Manual, Cold Spring Harbor, New York; Houston *et al.*, 1988, *Proc. Natl. Acad. Sci. USA* 85:5879-5883; and Bird *et al.*, 1988, *Science* 242:423-426.

Further provided by the present invention are anti-idiotypic antibodies (Ab2) and anti-anti-idiotypic antibodies (Ab3). See, e.g., M. Wettendorff et al., "Modulation of anti-tumor immunity by anti-idiotypic antibodies." In *Idiotypic Network and Diseases*, ed. by J. Cerny and J. Hiernaux, 1990 *J. Am. Soc. Microbiol.*, Washington DC: pp. 203-229]. These anti-idiotypic and anti-anti-idiotypic antibodies are produced using techniques well known to those of skill in the art. These antibodies may be used for a variety of purposes, including diagnostic and clinical methods and kits.

Under certain circumstances, it may be desirable to introduce a detectable label or a tag onto a SA18 gene product, antibody or other construct of the invention. As used herein, a detectable label is a molecule which is capable, alone or upon interaction with another molecule, of providing a detectable signal. Most desirably, the label is detectable visually, e.g. by fluorescence, for ready use in immunohistochemical analyses or immunofluorescent microscopy. For example, suitable labels include fluorescein isothiocyanate (FITC), phycoerythrin (PE), allophycocyanin (APC), coriphosphine-O (CPO) or tandem dyes, PE-cyanin-5 (PC5), and PE-Texas Red (ECD). All of these fluorescent dyes are commercially available, and their uses known to the art. Other useful labels include a colloidal gold label. Still other useful labels include radioactive compounds or elements. Additionally, labels include a variety of enzyme systems that operate to reveal a colorimetric signal in an assay, e.g., glucose oxidase (which uses glucose as a substrate) releases peroxide as a product which in the presence of peroxidase and a hydrogen donor such as tetramethyl benzidine (TMB) produces an oxidized TMB that is seen as a blue color. Other examples include horseradish peroxidase (HRP) or alkaline phosphatase (AP), and hexokinase in conjunction with glucose-6-phosphate dehydrogenase which reacts with ATP, glucose, and NAD⁺ to yield, among other products, NADH that is detected as increased absorbance at 340 nm wavelength.

Other label systems that are utilized in the methods of this invention are detectable by other means, e.g., colored latex microparticles [Bangs Laboratories, Indiana] in which a dye is embedded are used in place of enzymes to form conjugates with the target sequences provide a visual signal indicative of the presence of the resulting complex in applicable assays.

Methods for coupling or associating the label with a desired molecule are similarly conventional and known to those of skill in the art. Known methods of

label attachment are described [see, for example, Handbook of Fluorescent probes and Research Chemicals, 6th Ed., R. P. M. Haugland, Molecular Probes, Inc., Eugene, OR, 1996; Pierce Catalog and Handbook, Life Science and Analytical Research Products, Pierce Chemical Company, Rockford, IL, 1994/1995]. Thus,
5 selection of the label and coupling methods do not limit this invention.

The sequences, proteins, and fragments of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Suitable production techniques are well known to those of skill in the art. See, e.g., Sambrook et al, Molecular Cloning: A
10 Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternatively, peptides can also be synthesized by the well known solid phase peptide synthesis methods (Merrifield, *J. Am. Chem. Soc.*, 85:2149 (1962); Stewart and Young, Solid Phase Peptide Synthesis (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of
15 skill in the art and are not a limitation of the present invention.

In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for *in vitro*, *ex vivo* or *in vivo* delivery of therapeutic and immunogenic molecules. For example, in one embodiment, the simian Ad capsid
20 proteins and other simian adenovirus proteins described herein are used for non-viral, protein-based delivery of genes, proteins, and other desirable diagnostic, therapeutic and immunogenic molecules. In one such embodiment, a protein of the invention is linked, directly or indirectly, to a molecule for targeting to cells with a receptor for adenoviruses. Preferably, a capsid protein such as a hexon, penton, fiber or a fragment
25 thereof having a ligand for a cell surface receptor is selected for such targeting. Suitable molecules for delivery are selected from among the therapeutic molecules described herein and their gene products. A variety of linkers including, lipids, polyLys, and the like may be utilized as linkers. For example, the simian penton protein may be readily utilized for such a purpose by production of a fusion protein
30 using the simian penton sequences in a manner analogous to that described in Medina-Kauwe LK, et al, *Gene Ther.* 2001 May; 8(10):795-803 and Medina-Kauwe LK, et al, *Gene Ther.* 2001 Dec; 8(23): 1753-1761. Alternatively, the amino acid sequences of simian Ad protein IX may be utilized for targeting vectors to a cell

surface receptor, as described in US Patent Appln 20010047081. Suitable ligands include a CD40 antigen, an RGD-containing or polylysine-containing sequence, and the like. Still other simian Ad proteins, including, *e.g.*, the hexon protein and/or the fiber protein, may be used for these and similar purposes.

- 5, Still other adenoviral proteins of the invention may be used as alone, or in combination with other adenoviral protein, for a variety of purposes which will be readily apparent to one of skill in the art. In addition, still other uses for the adenoviral proteins of the invention will be readily apparent to one of skill in the art.

10 The compositions of this invention include vectors that deliver a heterologous molecule to cells, either for therapeutic or vaccine purposes. Such vectors, containing simian adenovirus DNA of SA18 and a minigene, can be constructed using techniques such as those described herein for the chimeric adenoviruses and such techniques as are known in the art. Alternatively, SA19 may be a source for sequences of the chimeric adenoviruses are described herein.

15 The following examples illustrate construction and use of several chimeric viruses, including Pan5/C1, hu5/Pan7 and hu5/SV25, and Pan6/Pan7. However, these chimera are illustrative only and are not intended to limit the invention to those illustrated embodiments.

20 **Example 1 - Construction of Pan5/C1 Chimeric Simian Viruses**

Five different adenoviruses initially isolated from the chimpanzee, AdC68 [US Patent 6,083,716], AdPan5, AdPan7, AdPan6 and AdC1 [US Patent 6,083,716] have been sequenced. See, International Application No. PCT/US02/33645, filed
25 November 2002 for the sequences of Pan5 [SEQ ID NO:1], Pan7 [SEQ ID NO:3], and Pan6 [SEQ ID NO:2]. This application also provides sequences for SV1, SV25 and SV39 [SEQ ID No. 4, 5, 6, respectively]. Sequence comparison of the capsid protein sequences predicted that AdC1 clearly belonged to a different serological subgroup than the other four chimpanzee derived adenoviruses.

30 However, attempts to cultivate AdC1 in HEK293 cells revealed it to be fastidious in its growth characteristics (data not shown) and therefore possibly unsuitable for use as a vector using the currently available E1 complementing cell lines. However, because of the obvious sequence dissimilarity of AdC1 capsid

protein sequence from the other chimpanzee derived adenoviruses (as well as the huAd5), chimeric adenovirus vectors were generated with the capsid characteristics of AdC1. In view of the above-mentioned drawbacks associated with only making hexon changes, more extensive replacements were made in the chimera described herein, *i.e.*, construction of chimeras where the replacement went beyond just the hexon, to achieve two goals. The first was to determine whether making extended replacements would allow for the rescue of viruses containing hexons of unrelated serotypes that may not otherwise be amenable to rescue. The second goal was to test whether the growth characteristics of adenovirus vectors such as AdPan5, that have been found in our laboratory to be able to be grown to high titer for the purpose of manufacture, would also be present in the chimeric virus, particularly when the hexon (and other capsid proteins) are derived from a virus such as AdC1 that are difficult to grow to a high yield in cell lines such as HEK293. An added bonus of extending the replacement to include the fiber protein would be to further increase the antigenic dissimilarity to beyond that afforded by a hexon change alone.

As an alternative to obtaining purified virus as source for adenoviral DNA to sequence, we have resorted to cloning restriction fragments of viral DNA obtained from infected cells ("Hirt prep"). The first adenovirus we have sequenced in this way is Simian Adenovirus. EcoRI digestion of the Simian Adenovirus yielded 7 fragments. Shotgun cloning yielded clones of the 5 internal fragments, which were cloned and sequenced. Completion of the sequencing was carried out by walking towards each of the ends of the genome. The map of the genome is shown in Figure 1.

A. *Construction of Two Pan5/C1 Chimeric Plasmids*

The overall approach towards constructing chimeric viruses was to first assemble the complete E1 deleted virus DNA into a single plasmid flanked by recognition sites for the restriction enzyme *Swa*I, digest the plasmid DNA with *Swa*I to release the virus DNA ends, and transfect the DNA into HEK293 cells to determine whether viable chimeric adenovirus could be rescued. Two chimeric virus plasmids were constructed, p5C1short and p5C1long.

The plasmid p5C1short harbors an E1 deleted Pan5 virus where an internal 15226 bp segment (18332 – 33557) has been replaced by a functionally analogous 14127 bp (18531 – 32657) from AdC1. This results in the replacement of

the Pan5 proteins hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from AdC1. The ClaI site at the left end of the AdC1 fragment is at the beginning of the hexon gene and the resulting protein is identical to the C1 hexon. The EcoRI site which constitutes the right end of the AdC1 fragment is within the E4 orf 7 part of the AdC1. The right end was ligated to a PCR generated right end fragment from AdPan5 such that the regenerated orf 7-translation product is chimeric between AdPan5 and AdC1.

The plasmid p5C1long harbors an E1 deleted Pan5 virus where an internal 25603 bp segment (7955 – 33557) has been replaced by a functionally analogous 24712 bp (7946 – 32657) from AdC1. This results in the replacement of the AdPan5 pre-terminal protein, 52/55 kD protein, penton base protein, protein VII, Mu, and protein VI with those from AdC1 in addition to those replaced in p5C1short. The AscI site at the left end of the AdC1 fragment is at the beginning of the DNA polymerase gene and results in a chimeric protein where the first 165 amino acids of the AdPan5 DNA polymerase has been replaced by a 167 amino acid segment from AdC1 DNA polymerase. In this N-terminal region, the homology between the AdPan5 and AdC1 DNA polymerase proteins is 81% (72% identity).

The plasmid pDVP5Mlu which contains the left end of AdPan5 was used as the starting plasmid for the chimeric vector construction.

The plasmid pDVP5Mlu was made as follows. A synthetic DNA fragment harboring recognition sites for the restriction enzymes SmaI, MluI, EcoRI and EcoRV respectively was ligated into pBR322 digested with EcoRI and NdeI so as to retain the origin of replication and the beta-lactamase gene. The left end of Pan5 extending to the MluI site (15135 bp) was cloned into this plasmid between the SmaI and MluI sites. The E1 gene was functionally deleted and replaced by a DNA fragment harboring recognition sites for the extremely rare cutter restriction enzyme sites I-CeuI and PI-SceI). The 2904 base pairs of the right end of Pan-5 was PCR amplified using the primers P5L [GCG CAC GCG TCT CTA TCG ATG AAT TCC ATT GGT GAT GGA CAT GC, SEQ ID NO:7] and P5ITR [GCG CAT TTA AAT CAT CAT CAA TAA TAT ACC TCA AAC, SEQ ID NO:8] using Tgo polymerase (Roche). The PCR product was cut with MluI and SwaI, and cloned between MluI and EcoRV of pDVP5Mlu to yield pPan5Mlu+RE. A 3193 bp fragment extending

from the MluI site (15135) to the ClaI (18328) site of Pan5 was then inserted between the same sites of pPan5Mlu+RE to yield pPan5Cla+RE. The 3671 bp ClaI (18531) to EcoRI (22202) fragment of the adenovirus C1 was cloned into pPan5Cla+RE between ClaI (16111) and EcoRI (16116) to yield pPan5C1delRI. The 10452 bp internal
5 EcoRI fragment of the adenovirus C1 (22202 – 32653) was cloned into the EcoRI site of pPan5C1delRI to yield p5C1short. To construct p5C1long, the AdC1 replacement was further extended by replacing the AscI – ClaI 10379 bp fragment of AdPan5 in p5C1short with the AdC1 AscI – ClaI 10591 bp fragment. Finally a green fluorescent protein (GFP) expression cassette was inserted into both p5C1short and p5C1long
10 between the I-CeuI and PI-SceI sites to yield p5C1shortGFP and p5C1longGFP respectively.

B. *Rescue of chimeric Pan5/C1 recombinant vector adenoviruses*

The plasmids p5C1shortGFP and p5C1longGFP were digested with the restriction enzyme SmaI and transfected into HEK 293 cells. A typical adenovirus
15 induced cytopathic effect was observed. The rescue of recombinant chimeric adenovirus from the p5C1longGFP transfection was confirmed by collecting the supernatant from the transfection and re-infecting fresh cells which were found to be transduced as determined by GFP expression. Viral DNA prepared from the chimeric recombinant virus was digested with several restriction enzymes and found to have
20 the expected pattern on electrophoresis (data not shown).

The chimeric adenoviral construct with the shorter replacement p5C1short encodes the C1 proteins hexon and fiber as well as the intervening open reading frames for endoprotease, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, and protein VIII. (The E3 region is also included within this region
25 but is unlikely to impact on the viability of the chimeric virus). When the replacement was extended to include the additional AdC1 proteins pTP (pre-terminal protein), 52/55 kDa protein, penton base, protein VII, Mu, and protein VI, there was no difficulty in rescuing viable chimeric virus. In this experiment, the chimeric adenovirus construction strategy utilized the presence of AscI and ClaI restriction
30 enzyme sites present on the genes for DNA polymerase and hexon respectively on both AdPan5 and AdC1.

The reasons for the relatively higher yield of the chimeric virus compared to the wild-type AdC1 virus are not clear. In the growth of the 5C1

chimeric virus in 293 cells, the adenoviral early region gene products of E1 and E4 are derived from Ad5 and AdPan5 respectively. The E1 and E4 gene products bind, regulate and de-repress several cellular transcription complexes and coordinate their activity towards viral multiplication. Thus it is possible that the E1 gene products supplied in *trans* from the 293 cells and the E4 gene products from AdPan5 are more optimal in the human 293 cell background than are the equivalent AdC1 gene products. This may also apply to the activity of the major late promoter whose activity is responsible for the transcription of the capsid protein genes. In the chimeric virus, the major late promoter, and the protein IVa2 which transactivates it, are derived from AdPan5. However the E2 gene products required for adenoviral DNA replication pTP and single-stranded DNA – binding protein are derived from AdC1. The adenoviral DNA polymerase, which complexes with pTP, is chimeric in Ad5C1 but mostly AdPan5 derived.

Example 2 - Construction of Ad5 Chimeric Simian Viruses

Plasmids have been constructed where the structural proteins derive from the chimpanzee adenovirus Pan 7 and the flanking sequences are derived from human Ad5 (the commonly used vector strain). The Adhu5-Pan7 chimeric adenovirus has been rescued, demonstrating that the chimeric virus construction method used to derive the chimeric virus is broadly applicable.

A. Construction of the Ad5 – Pan 7 chimeric adenovirus

A plasmid was constructed which harbors the complete (E1 deleted) chimeric genome in order to establish that the chimeric adenovirus is viable, and then transfected the plasmid into the E1 complementing cell line HEK 293. It was found that the recombinant virus could be rescued. The chimeric adenovirus genome that was constructed is composed of a left end segment derived from Ad5 that contributes the ITR, the E1 deletion region containing the transgene expression cassette, the pIX and IVa2 genes and 954 C-terminal amino acids of the polymerase gene (which is transcribed in the right to left direction from the bottom strand). Ad5 also contributes the right end of the chimeric genome containing the E4 genes and the right ITR. All the other genes present in the central part of the chimeric construct are derived from the chimpanzee adenovirus Pan 7 including the N-terminal 235 amino acids of a chimeric DNA polymerase.

In order to construct the plasmid which harbors the complete (E1 deleted) chimeric genome, the starting plasmid was pBRAd5lere which is comprised of three parts; the bacterial origin of replication and ampicillin resistance gene derived from the plasmid pBR322, the left end of an Ad5 derived E1 deleted vector
 5 extending from the left ITR to the StuI site located at base pair number 5782 of the wild-type Ad5 genome (the E1 deletion extends from base pair 342 to 3533 of the wild-type Ad5 genome), and the right end of Ad5 extending from the StuI site at base pair number 31954 of the wild-type Ad5 genome to the right end of the right ITR. The PacI sites located adjacent to the two ITRs are used to release the Ad5 genome
 10 from the bacterial plasmid backbone. The fragment containing I-CeuI and the PI-SceI sites which is located in place of the E1 deletion is used to insert transgene cassettes.

A synthetic DNA oligomer was inserted at the StuI site containing sites for AscI, XbaI and EcoRI, which allowed the creation of the plasmid pAd5endsAscRI where using PCR, the Ad5 polymerase gene was extended to base
 15 pair #8068 of the wild-type Ad5 genome and incorporating a newly created AscI site at this location by silent mutagenesis of the polymerase gene (translated from the bottom strand) as depicted below.

Original sequence

20 GCG ACG GGC CGA [SEQ ID NO:16]
 CGC TGC CCG GCT
 Arg Arg Ala Ser [SEQ ID NO:17]

Mutated sequence (The AscI recognition site is underlined)

25 GCG GCG CGC CGA [SEQ ID NO:18]
 CGC TGC CCG GCT
 Arg Arg Ala Ser [SEQ ID NO: 17]

The Pan 7 fiber containing region was amplified by PCR (mutating the fiber stop codon from TGA to TAA to provide a polyadenylation signal similar to that in Ad5) and inserted into the EcoRI site to yield pAd5endsP7fib. Several cloning steps led to the construction of pH5C7H5 where the complete chimeric adenoviral
 30 genome has been assembled. A transgene cassette expressing GFP (green fluorescent protein) was inserted between the I-CeuI and PI-SceI sites of pH5C7H5. The final construct was digested with PacI to separate the adenoviral genome from the plasmid

backbone and transfected into HEK 293 cells. The cell lysate was harvested 2 weeks later, and the chimeric adenovirus was amplified and purified by standard methods.

5 B. Construction of the Ad5 – Simian virus 25 (SV-25) chimeric adenovirus

[N.B. Simian virus 25 (ATCC catalog number VR-201) is distinct from the chimpanzee adenovirus Simian adenovirus 25 ATCC catalog number VR-594]

10 The construction of the Ad5 based chimeric adenovirus where the left and right end segments are derived from Ad5 and the central portion was derived from the monkey adenovirus SV-25 was carried out in a manner completely analogous to that described above for the chimeric adenovirus described above that is chimeric between Ad5 and the chimpanzee adenovirus Pan 7. Thus, the chimeric adenovirus genome that was constructed is composed of a left end segment derived
15 from Ad5 that contributes the ITR, the E1 deletion region containing the transgene expression cassette, the pIX and IVa2 genes and 956 C-terminal amino acids of the polymerase gene. Ad5 also contributes the right end of the chimeric genome containing the E4 genes and the right ITR. [Additionally, the left end of the Ad5 genome was extended beyond that present in pH5C7H5 so that 454 base pairs of the
20 Ad5 left end was present. Although not absolutely essential, this was done in order to improve packaging efficiency.] All the genes present in the central part of the chimeric construct are derived from the monkey adenovirus SV-25 including the N-terminal 230 amino acids of a chimeric DNA polymerase. The starting plasmid for the construction of the chimeric genome was pAd5endsAscRI which contains both
25 the left and right ends of Ad5 as well as the created (by silent mutation) AscI site in the polymerase gene where Ad5-SV25 chimeric fusion was made (as was done for the Ad5 – Pan 7 chimeric adenovirus). In the final construct pH5S25H5, the SV25 genome segment has been incorporated by sequential cloning steps, including creation of an AscI site at the ligation junction within the polymerase coding
30 sequence. A transgene cassette expressing GFP (green fluorescent protein) was inserted between the I-CeuI and PI-SceI sites of pH5S25H5. The final construct was digested with PacI to separate the adenoviral genome from the plasmid backbone and

transfected into HEK 293 cells. The cell lysate was harvested 2 weeks later, and the chimeric adenovirus was amplified and purified by standard methods.

Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus. The portion of the genome replaced by DNA from Pan7 is indicated.

5

Example 3 – Pan5 - C1 chimeric vector of invention as a delivery vehicle for immunogenic compositions

A Pan 5 (Simian adenovirus 22, a subgroup E adenovirus, also termed C5) - C1 (Simian adenovirus 21, a subgroup B adenovirus) chimeric expressing the Ebola virus (Zaire) glycoprotein (C5C1C5-CMVGP) was constructed as a model antigen in order to test the efficacy of the vector C5C1C5-CMVGP as a vaccine; this vector has been compared it to the Adhu5 based vector (H5-CMVGP). Compared to H5-CMVGP, the C5C1-CMVGP vector yielded only a slightly decreased level of GP expression in transduced A549 cells.

15 Thereafter, GP-specific T cell and B cell responses elicited in B10BR mice vaccinated intramuscularly with either 5×10^{10} H5-CMVGP or C5C1-CMVGP vectors were compared.

The C5C1C5-CMVGP vector appeared to induce lower frequencies of gamma interferon producing CD8+ T cells with kinetics slower than the H5-CMVGP vector as determined by intracellular cytokine staining using a H-2k restricted GP-specific peptide as stimulant. The total IgG response to GP, measured by ELISA, was equivalent in serum from mice vaccinated with the C5C1C5-CMVGP or the H5-CMVGP vectors. However, the C5C1C5-CMVGP vector induced a more potent Th1 type response while the H5-CMVGP vector stimulated a more balanced Th1/Th2 type response. In a survival study, mice were vaccinated as above and challenged 28 days later with 200 LD/50 mouse-adapted Ebola Zaire virus. 100% survival was seen for both groups.

Example 4 – Generation of Chimeric Pan6/Pan7 Vectors

30 A panel of GFP expressing vectors were generated. This panel includes vectors that are chimeric between Pan 6 and Pan 7 where (a) the hexon protein of Pan 7 was replaced by that of Pan 6 (termed C767), (b), the fiber protein of Pan7 was

replaced by that of Pan 6 (termed C776), (c) both the hexon and fiber proteins of the Pan 7 vector have been replaced by those from Pan 6 (termed C766).

5 The chimeric virus termed C767 was constructed essentially as described above for the C5C1C5 virus in Example 1. However, due to substantial homology between the Pan6 and Pan7 sequences 5' to the hexon sequence, it was not necessary to substitute the 5' end of the genome between the penton and the pol gene.

The chimeric vector C767 was compared to the C776, C766, the parent C6, and the parent C7, each expressing GFP.

10 Balb/C mice (25 per group) were immunized intramuscularly with either Pan 6 or Pan 7 (10^{10} particles). Re-administration (10^{11} particles i.v., by tail vein injection) was attempted 3 weeks later using each of the five GFP expressing vectors (C6-GFP, C7-GFP, and the three chimeric vectors). Three days later the level of liver transduction was estimated qualitatively by examining liver sections for the presence of GFP expression and quantitatively by estimating copies of GFP DNA by Taqman
15 analysis. Administration of either one of the two chimpanzee adenovirus vectors does not affect the transduction efficiency of the other vector, while re-administration of the same vector is severely compromised. The data showed that antibodies to both hexon and fiber are important in preventing re-administration of adenoviral vectors.

20

All publications cited in this specification are incorporated herein by reference, as are the priority documents, US Patent Application 60/575,429, filed March 28, 2004; US Patent Application No. 60/566,212, filed April 28, 2004, and US Patent Application No. 10/465,302, filed June 20, 2003. While the invention has been
25 described with reference to a particularly preferred embodiment, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope of the appended claims.

What is claimed is:

1. A method of efficiently culturing a chimeric adenovirus in a selected host cell, said chimeric adenovirus being from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the steps of:

(a) generating a chimeric adenovirus comprising:

(i) adenovirus sequences of the left terminal end and right terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the 5' inverted terminal repeat (ITRs), and said right end region comprising the 3' inverted terminal repeat (ITRs); and

(ii) the internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the late genes encoding the penton, hexon, and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', a left terminal region of the first adenovirus, the internal region of the parental adenovirus, and the right terminal region of the first adenovirus; and

b) culturing said chimeric adenovirus in the presence of functional adenovirus E1a, E1b, and E4 ORF6 genes from the first adenovirus or from an adenovirus serotype which transcomplements the first adenovirus, and further in the presence of necessary adenoviral structural genes from the left end of the adenovirus.

2. The method according to claim 1, wherein the internal region of the parental adenovirus further comprises one or more functional adenovirus genes selected from the group consisting of Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and protein VI.

3. The method according to claim 1, wherein the polymerase, terminal protein and 52/55 kDa protein functions are provided in *trans*.

4. The method according to claim 1, wherein the first adenovirus further comprises the polymerase, terminal protein and 52/55 kDa protein functions.
5. The method according to claim 1, wherein the chimeric adenovirus comprises the adenoviral late genes 1, 2, 3, 4, and 5 of the parental adenovirus.
6. The method according to claim 1, wherein the selected host cell stably contains one or more of the adenovirus E1a, E1b or E4 ORF6 functions.
7. The method according to claim 1, wherein the chimeric adenovirus comprises one or more of the adenovirus E1a, E1b or E4 ORF6 of the first adenovirus.
8. The method according to claim 1, wherein the first adenovirus is of human origin.
9. The method according to claim 1, wherein the first adenovirus is of simian origin.
10. The method according to claim 1, further comprising the step of isolating the chimeric adenovirus.
11. A method for generating a chimeric adenovirus for growth in a selected host cell, said chimeric adenovirus being derived from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the step of generating a chimeric adenovirus comprising:
 - 5' and 3' terminal regions of a first adenovirus which grows in a selected host cell type, said 5' terminal regions comprising the 5' inverted terminal repeat (ITRs) and necessary E1 gene functions, and said 3' terminal regions comprising inverted terminal repeat (ITRs) and necessary E4 gene functions; and
 - internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the hexon, penton base and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', the 5' terminal region of the first adenovirus, the internal region of the parental adenovirus, and the 3' terminal regions of the first adenovirus.

12. A chimeric adenovirus produced according to the method of claim 1.

13. A chimeric adenovirus comprising a hexon protein of a selected adenovirus serotype which is incapable of efficient growth in a selected host cell, said modified adenovirus comprising:

(a) adenovirus sequences of the left terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the E1a, E1b and 5' inverted terminal repeat (ITRs);

(b) adenovirus sequences of the internal region of the selected adenovirus serotype which is incapable of efficient growth in the selected host cell, said internal region comprising the genes encoding the penton, hexon and fiber of the selected adenovirus;

(c) adenovirus sequences of the right terminal end of the first adenovirus, said right end region comprising the necessary E4 gene functions and the 3' inverted terminal repeat (ITRs),

wherein the resulting chimeric adenovirus comprises adenoviral structural and regulatory proteins necessary for infection and replication.

14. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus further comprises the IIIa, 52/55kDa and terminal protein (pTP) of the selected adenovirus serotype.

15. The chimeric adenovirus according to claim 13, wherein chimeric adenovirus comprises the polymerase of the first adenovirus.

16. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus expresses a functional chimeric protein formed from the first adenovirus and the selected adenovirus, said chimeric protein is selected from the group consisting of polymerase, terminal protein, 52/55 kDa protein, and IIIa.

17. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus comprises the terminal protein, 52/55 kDa, and/or IIIa of the selected adenovirus.

18. A host cell comprising a chimeric adenovirus according to claim 12.

19. The host cell according to claim 18, wherein said host cell is a human cell.

20. An isolated simian adenovirus nucleic acid sequence selected from the group consisting of:

(a) SA18 having the sequence of nucleic acids 1 to 31967 of SEQ ID NO:12 and

(b) a nucleic acid sequence complementary to the sequence of any of (a) to (f).

21. An isolated simian adenovirus serotype nucleic acid sequence selected from one or more of the group consisting of:

(a) 5' inverted terminal repeat (ITR) sequences;

(b) the adenovirus E1a region, or a fragment thereof selected from among the 13S, 12S and 9S regions;

(c) the adenovirus E1b region, or a fragment thereof selected from among the group consisting of the small T, large T, IX, and IVa2 regions;

(d) the E2b region;

(e) the L1 region, or a fragment thereof selected from among the group consisting of the 28.1 kD protein, polymerase, agnoprotein, 52/55 kD protein, and IIIa protein;

(f) the L2 region, or a fragment thereof selected from the group consisting of the penton, VII, VI, and Mu proteins;

(g) the L3 region, or a fragment thereof selected from the group consisting of the VI, hexon, or endoprotease;

(h) the 2a protein;

- (i) the L4 region, or a fragment thereof selected from the group consisting of the 100 kD protein, the 33 kD homolog, and VIII;
- (j) the E3 region, or a fragment thereof selected from the group consisting of E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, and E3 ORF9;
- (k) the L5 region, or a fragment thereof selected from a fiber protein;
- (l) the E4 region, or a fragment thereof selected from the group consisting of E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, and E4 ORF1; and
- (m) the 3' ITR, of any of SA18 SEQ ID NO:12, or a sequence complementary to any of (a) to (m).

22. A simian adenovirus protein encoded by the nucleic acid sequence according to claim 21.

23. A composition comprising a simian adenovirus capsid protein according to claim 22 linked to a heterologous molecule for delivery to a selected host cell.

24. A method for targeting a cell having an adenoviral receptor comprising delivering to a subject a composition according to claim 23.

25. A nucleic acid molecule comprising a heterologous simian adenoviral sequence according to claim 21.

26. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence encodes an adenoviral gene product and is operatively linked to regulatory control sequences which direct expression of the adenoviral gene product in a host cells.

27. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence comprises the E1a region of SA18 SEQ ID NO:12.

28. A pharmaceutical composition comprising the nucleic acid molecule according to claim 27 and a physiologically compatible carrier.
29. A recombinant adenovirus having a capsid comprising a protein selected from the group consisting of:
- (a) a hexon protein of SA18, SEQ ID NO 13, or a unique fragment thereof;
 - (b) a penton protein of SA18, SEQ ID NO: 14, or a unique fragment thereof;
 - (c) a fiber protein of SA18, SEQ ID NO: 15, or a unique fragment thereof.
30. The recombinant adenovirus according to claim 29, wherein the capsid is of an artificial serotype.
31. The recombinant adenovirus according to claim 29, wherein said virus further comprises a heterologous gene operatively linked to sequences which direct expression of said gene in a host cell.
32. The recombinant adenovirus according to claim 29, further comprising 5' and 3' adenovirus cis-elements necessary for replication and encapsidation.
33. The recombinant adenovirus according to claim 29, wherein said vector lacks all or a part of the E1 gene.
34. A host cell comprising a heterologous nucleic acid molecule comprising the nucleic acid sequence according to claim 21.
35. The host cell according to claim 34, wherein said host cell is stably transformed with the nucleic acid molecule.

36. The host cell according to claim 34, wherein said host cell expresses one or more adenoviral gene products from said nucleic acid molecule, said adenoviral gene products selected from the group consisting of E1a, E1b, E2a, and E4 ORF6.

37. The host cell according to claim 34, wherein said host cell is stably transformed with a nucleic acid molecule comprising the simian adenovirus inverted terminal repeats.

38. A composition comprising a recombinant virus according to claim 29 in a pharmaceutically acceptable carrier.

39. A method for delivering a heterologous gene to a mammalian cell comprising introducing into said cell an effective amount of the recombinant virus according to claim 29.

40. A method for repeat administration of a heterologous gene to a mammal comprising the steps of:

- (a) introducing into said mammal a first vector which comprises the heterologous gene and
 - (b) introducing into said mammal a second vector which comprises the heterologous gene;
- wherein at least the first virus or the second vector is a virus according to claim 29 and wherein the first and second recombinant vector are different.

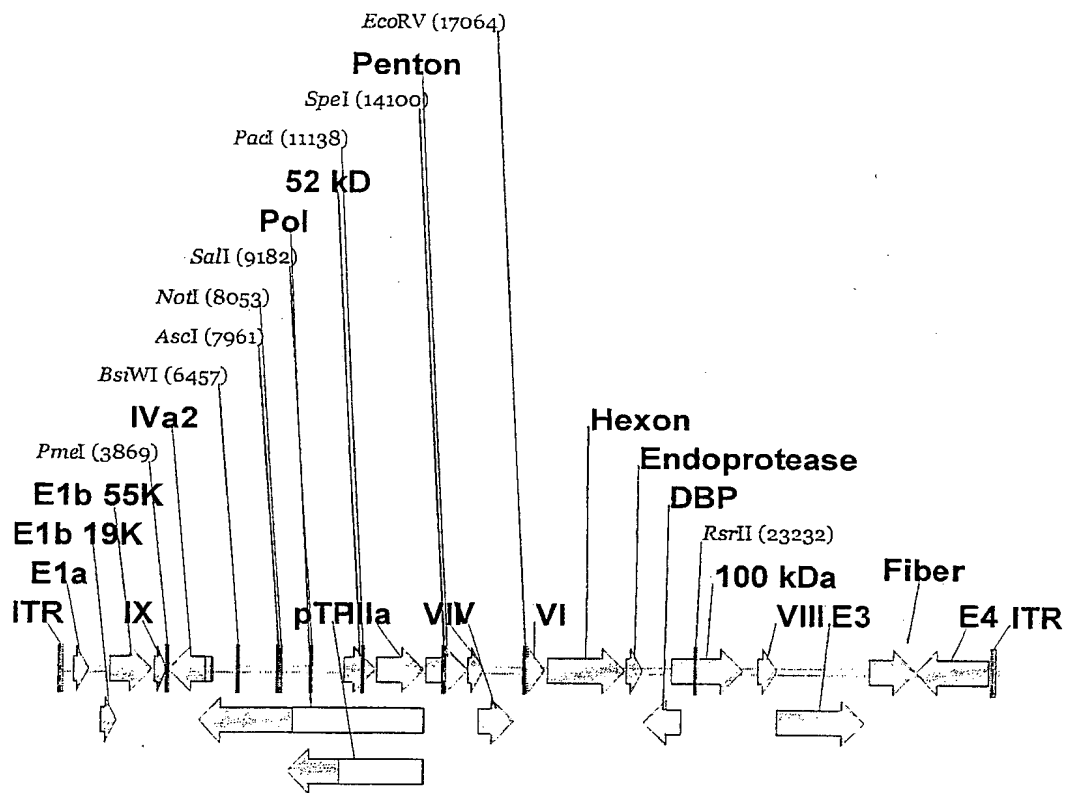
41. A method for producing a selected gene product comprising infecting a mammalian cell with the recombinant virus according to claim 29, culturing said cell under suitable conditions and recovering from said cell culture the expressed gene product.

42. A method for eliciting an immune response in a mammalian host against an infective agent comprising administering to said host an effective amount of the

recombinant adenovirus of claim 29, wherein said heterologous gene encodes an antigen of the infective agent.

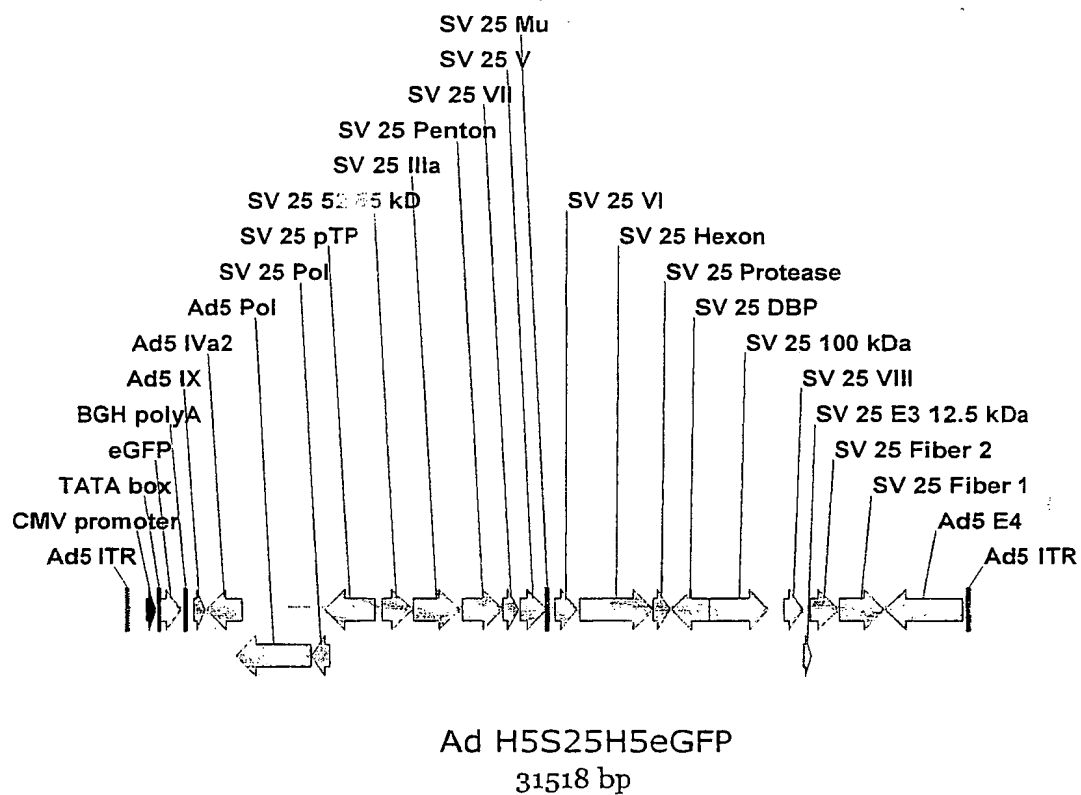
43. The method according to claim 42, comprising the step of priming the host with a DNA vaccine comprising the heterologous gene prior to administering the recombinant adenovirus.

Fig. 1



Simian Adenovirus
34302 bp

Fig. 2



SEQUENCE LISTING

<110> The Trustees of the University of Pennsylvania
Roy, Soumitra
Wilson, James M.

<120> Methods of Generating Chimeric Adenoviruses and Uses For Such
Chimeric Adenoviruses

<130> UPN-P3067PCT

<150> US 10/465,302
<151> 2003-06-20

<150> US 60/566,212
<151> 2004-04-28

<150> US 60/575,429
<151> 2004-05-28

<160> 18

<170> PatentIn version 3.2

<210> 1
<211> 36462
<212> DNA
<213> chimpanzee adenovirus serotype Pan5

<400> 1
catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg aggtatttga 60
atttgggggat gcggggcggt gattggctgc gggagcggcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgccgag 360
ggccgagtag actttgaccg attacgtggg ggtttctgatt accgtatttt tcacctaaat 420
ttccgcgtac ggtgtcaaag tccggtgttt ttacgtagggt gtcagctgat cgccagggta 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
gtaatgtttt cctggctact gggaacgaga ttctggaact ggtggtggac gccatgatgg 660
gtgacgacct tccggagccc cctaccccat ttgaagcgcc ttcgctgtac gatttgatatg 720
atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
gcgatgccgc gctgctggct gccgagcagg ctaatacggg ctctggctca gacagcgatt 840
cctctctcca taccgagaga cccggcagag gtgagaaaaa gatcccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg 960
aggaggcgat tcgagctgca gcgaaccagg gagtgaaaac agcgagcgag ggcttttagcc 1020
tggaactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata 1080
ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
acagtaagtg tgattaactt tagctgggga ggcagagggt gactgggtgc tgactggttt 1200

atttatgtat atgtttttta tgtttaggtc ccgctctctga cgtagatgag acccccacta	1260
cagagtgcac ttcatcaccc ccagaaattg gcgaggaacc gcccgaagat attattcata	1320
gaccagttgc agtgagagtc accgggcgta gagcagctgt ggagagtttg gatgacttgc	1380
tacagggttg ggatgaacct ttggacttgt gtaccgggaa acgccccagg cactaagtgc	1440
cacacatgtg tgtttactta aggtgatgtc agtattttata ggggtgtggag tgcaataaaa	1500
tccgtgttga ctttaagtgc gtggtttatg actcaggggt ggggactgtg ggtatataag	1560
caggtgcaga cctgtgtggg cagttcagag caggactcat ggagatcttg acagtcttgg	1620
aagactttca ccagactaga cagctgctag agaactcatc ggagggagtc tcttacctgt	1680
ggagattctg cttcgggtggg cctctagcta agctagtcta tagggccaag caggattata	1740
aggatcaatt tgaggatatt ttgagagagt gtcctggtat ttttgactct ctcaacttgg	1800
gccatcagtc tcactttaac cagagtattc tgagagccct tgacttttct actcctggca	1860
gaactaccgc cgcggtagcc ttttttgcct ttatccttga caaatggagt caagaaaccc	1920
atttcagcag ggattaccgt ctggactgct tagcagtagc tttgtggaga acatggaggt	1980
gccagcgcct gaatgcaatc tccggctact tgccagtaca gccggtagac acgctgagga	2040
tcctgagtct ccagtcaccc caggaacacc aacgccgcca gcagccgcag caggagcagc	2100
agcaagagga ggaccgagaa gagaacctga gagccggtct ggaccctccg gtggcggagg	2160
aggaggagta gctgacttgt ttcccagct ggcgccgggtg ctgactaggt cttccagtgg	2220
acgggagagg gggattaagc gggagaggca tgaggagact agccacagaa ctgaactgac	2280
tgtcagtctg atgagtcgca ggcgccaga atcgggtgtg tggcatgagg tgcagtcgca	2340
ggggatagat gaggtctcag tgatgcatga gaaatattcc ctagaacaag tcaagacttg	2400
ttggttggag cccgaggatg attgggaggt agccatcagg aattatgcca agctggctct	2460
gaggccagac aagaagtaca agattaccaa actgattaat atcagaaatt cctgctacat	2520
ttcagggaat ggggcccagg tggagatcag taccaggag aggggtggcct tcagatgctg	2580
catgatgaat atgtaccggg ggggtggtgg catggaggga gtcaccttta tgaacgcgag	2640
gttcaggggt gatgggtata atgggggtgt ctttatggcc aacaccaagc tgacagtgca	2700
cggatgctcc ttctttggct tcaataacat gtgcattgag gcctggggca gtgtttcagt	2760
gaggggatgc agtttttcag ccaactggat gggggctcgtg ggcagaacca agagcatggt	2820
gtcagtgaag aaatgcctgt tcgagagggt ccacctgggg gtgatgagcg agggcgaagc	2880
caaagtcaaa cactgcgctt ctaccgagac gggctgcttt gtactgatca agggcaatgc	2940
caaagtcaag cataatatga tctgtggggc ctcggatgag cgcggctacc agatgctgac	3000
ctgcgccggt gggaaacagcc atatgctagc caccgtgcat gtggcctcgc acccccgcaa	3060
gacatggccc gagttcgagc acaacgtcat gaccgctgc aatgtgcacc tggggtcccg	3120
ccgaggcatg ttcatgccct accagtgcaa catgcaattt gtgaagggtg tgctggagcc	3180
cgatgccatg tccagagtga gcctgacggg ggtgtttgac atgaatgtgg agctgtggaa	3240

aattctgaga	tatgatgaat	ccaagaccag	gtgccgggccc	tgcgaatgcg	gaggcaagca	3300
cgccaggctt	cagcccgtgt	gtgtggaggt	gacggaggac	ctgcgacccg	atcatttggt	3360
gttgtcctgc	aacgggacgg	agttcggctc	cagcggggaa	gaatctgact	agagtgagta	3420
gtgtttggga	ctgggtggga	gcctgcatga	tgggcagaat	gactaaaatc	tgtgtttttc	3480
tgcgcagcag	catgagcggg	agcgcctcct	ttgagggagg	ggtattcagc	ccttatctga	3540
cggggcgtct	cccctcctgg	gcgggagtg	gtcagaatgt	gatgggatcc	acggtggacg	3600
gccggcccgt	gcagcccgcg	aactcttcaa	ccctgacctt	cgcgaccctg	agctcctcgt	3660
ccgtggacgc	agctgccgcc	gcagctgctg	cttccgccgc	cagcgccgtg	cgcggaatgg	3720
ccctggggcg	cggctactac	agctctctgg	tggccaactc	gagttccacc	aataatcccc	3780
ccagcctgaa	cgaggagaag	ctgctgctgc	tgatggccca	gctcgaggcc	ctgaccacgc	3840
gcctggggcg	gctgaccag	caggtggctc	agctgcaggc	ggagacgcgg	gccgcggttg	3900
ccacggtgaa	aaccaaataa	aaaatgaatc	aataaataaa	cggagacggt	tgttgatttt	3960
aacacagagt	cttgaatctt	tatttgattt	ttcgcgcgcg	gtaggccctg	gaccaccggt	4020
ctcgatcatt	gagcaccgcg	tggatctttt	ccaggacccg	gtagagggtg	gcttgatgtg	4080
tgaggtagat	gggcatgagc	ccgtcccggg	ggtggaggta	gctccattgc	agggcctcgt	4140
gctcgggggt	ggtgttgtaa	atcaccaggt	catagcaggg	gcgcagggcg	tgggtgtgca	4200
cgatgtcctt	gaggaggaga	ctgatggcca	cgggcagccc	cttgggtgtag	gtgttgacga	4260
acctgttgag	ctgggagggg	tgcattgcgg	gggagatgag	atgcatcttg	gcctggatct	4320
tgagattggc	gatgttccc	cccagatccc	gccggggggt	catgttggtg	aggaccacca	4380
gcacggtgta	tccggtgcac	ttggggaatt	tgtcatgcaa	cttgggaagg	aaggcgtgaa	4440
agaatttgga	gacgcccttg	tgaccgcca	ggttttccat	gcactcatcc	atgatgatgg	4500
cgatgggccc	gtgggcggcg	gcttgggcaa	agacgtttcg	ggggtcggac	acatcgtagt	4560
tgtggtcctg	ggtgagctcg	tcataggcca	ttttaatgaa	tttggggcg	aggggtgccc	4620
actgggggac	gaaggtgccc	tcgatcccgg	gggcgtagtt	gccctcgcag	atctgcatct	4680
cccaggcctt	gagctcggag	ggggggatca	tgtccacctg	cggggcgatg	aaaaaacgg	4740
tttccggggc	gggggagatg	agctggggcg	aaagcagggt	ccggagcagc	tgggacttgc	4800
cgcagccggt	ggggccgtag	atgaccccga	tgaccggctg	caggtggtag	ttgagggaga	4860
gacagctgcc	gtcctcgcgg	aggagggggg	ccacctcggt	catcatctcg	cgcacatgca	4920
tgttctcgcg	cacgagttcc	gccaggaggc	gctcgccccc	aagcgagagg	agctcttgca	4980
gcgaggcgaa	gtttttcagc	ggcttgagcc	cgctcgcccat	gggcattttg	gagaggggtct	5040
gttgcaagag	ttccagacgg	tcccagagct	cggtgatgtg	ctctagggca	tctcgatcca	5100
gcagacctcc	tcgtttcgcg	ggttggggcg	actgcgggag	tagggcacca	ggcgatgggc	5160
gtccagcgag	gccagggtcc	ggtccttcca	ggggcgaggg	gtccgcgtca	gcgtggtctc	5220
cgtcacggtg	aaggggtg	cgccgggctg	ggcgcttg	aggggtgcgt	tcaggctcat	5280
ccggctggtc	gagaaccgct	cccggctcgg	gccctgcgcg	tcggccagggt	agcaattgag	5340

catgagttcg	tagttgagcg	cctcggccgc	gtggcccttg	gcgcggagct	tacctttgga	5400
agtgtgtccg	cagacgggac	agaggagggg	cttgagggcg	tagagcttgg	gggcgaggaa	5460
gacggactcg	ggggcgtagg	cgtccgcgcc	gcagctggcg	cagacggtct	cgcactccac	5520
gagccaggtg	aggtctggcc	ggtcggggtc	aaaaacgagg	tttcctccgt	gctttttgat	5580
gcgtttctta	cctctggtct	ccatgagctc	gtgtccccgc	tgggtgacaa	agaggctgtc	5640
cgtgtccccg	tagaccgact	ttatgggccg	gtcctcgagc	ggggtgccgc	ggtcctcgtc	5700
gtagaggaa	cccgccact	ccgagacgaa	ggcccgggtc	caggccagca	cgaaggaggc	5760
cacgtgggag	gggtagcgg	cgttgtccac	cagcgggtcc	accttctcca	gggtatgcaa	5820
gcacatgtcc	ccctcgtcca	catccaggaa	ggtgattggc	ttgtaagtgt	agggcacgtg	5880
accgggggtc	ccggccgggg	gggtataaaa	gggggcgggc	ccctgctcgt	cctcactgtc	5940
ttccggatcg	ctgtccagga	gcgccagctg	ttggggtagg	tattccctct	cgaaggcggg	6000
catgacctcg	gcactcaggt	tgtcagtttc	tagaaacgag	gaggatttga	tattgacggt	6060
gccgttgag	acgcctttca	tgagcccctc	gtccatctgg	tcagaaaaga	cgatcttttt	6120
gttgtcgagc	ttggtggcga	aggagccgta	gagggcgttg	gagagcagct	tggcgatgga	6180
gcgcatggtc	tggttctttt	ccttgtcggc	gcgctccttg	gcggcgatgt	tgagctgcac	6240
gtactcgcgc	gccacgcact	tccattcggg	gaagacggtg	gtgagcttgt	cgggcacgat	6300
tctgaccgcg	cagccgcggt	tgtgcagggt	gatgagggtc	acgctggtgg	ccacctcgcc	6360
gcgcaggggc	tcgttggtcc	agcagaggcg	cccgcccttg	cgcgagcaga	aggggggcag	6420
cgggtccagc	atgagctcgt	cgggggggtc	ggcgtccacg	gtgaagatgc	cgggcaggag	6480
ctcggggtcg	aagtagctga	tgcagggtgc	cagatcgtcc	agcgccgctt	gccagtcgcg	6540
cacggccagc	gcgcgctcgt	aggggctgag	gggcgtgccc	cagggcatgg	ggtgctgag	6600
cgcggaggcg	tacatgccgc	agatgtcgta	gacgtagagg	ggctcctcga	ggacgccgat	6660
gtaggtgggg	tagcagcgcc	ccccgcggat	gctggcgcg	acgtagtcgt	acagctcgtg	6720
cgagggcgcg	aggagcccgg	tgccgagggt	ggagcgctgc	ggcttttcgg	cgcggtagac	6780
gatctggcgg	aagatggcgt	gggagttgga	ggagatggtg	ggcctctgga	agatgttgaa	6840
gtgggcgtgg	ggcagtcgga	ccgagtcctt	gatgaagtgg	gcgtaggagt	cctgcagctt	6900
ggcgacgagc	tcggcggtga	cgaggacgtc	cagggcgag	tagtcgaggg	tctcttggtg	6960
gatgtcgtag	ttgagctggc	ccttctgctt	ccacagctcg	cggttgagaa	ggaactcttc	7020
gcggtccttc	cagtactctt	cgagggggaa	cccgtcctga	tcggcacggt	aagagcccac	7080
catgtagaac	tggttgacgg	ccttgtaggc	gcagcagccc	ttctccacgg	ggagggcgta	7140
agcttgcgcg	gccttgcgca	gggagggtgtg	ggtgagggcg	aagggtgtcg	gcaccatgac	7200
cttgaggaac	tggtgcttga	agtcgaggtc	gtcgcagccg	ccctgtctcc	agagctggaa	7260
gtccgtgcgc	ttcttgtagg	cgggggttggg	caaagcgaaa	gtaacatcgt	tgaagaggat	7320
cttgcccgcg	cggggcatga	agttgcgagt	gatgcggaaa	ggctggggca	cctcggccccg	7380

gttggtgatg	acctgggagg	cgaggacgat	ctcgtcgaag	ccgttgatgt	tgtgcccagc	7440
gatgtagagt	tccacgaatc	gcggggcgcc	cttgacgtgg	ggcagcttct	tgagctcgtc	7500
gtaggtgagc	tcggcggggt	cgctgaggcc	gtgctgctcg	agggcccagt	cggcgaggtg	7560
ggggttgccg	ccgaggaagg	aagtccagag	atccacggcc	agggcggctc	gcaagcggtc	7620
ccggtactga	cggaactgct	ggcccacggc	catttttttcg	ggggtgacgc	agtagaaggt	7680
gcgggggtcg	ccgtgccagc	ggtcccactt	gagctggagg	gcgaggtcgt	gggcgagctc	7740
gacgagcggc	gggtccccgg	agagtttcat	gaccagcatg	aaggggacga	gctgcttgcc	7800
gaaggacccc	atccaggtgt	aggttttcac	gtcgtaggtg	aggaagagcc	tttcggtgcg	7860
aggatgagag	ccgatgggga	agaactggat	ctcctgccac	cagttggagg	aatggctgtt	7920
gatgtgatgg	aagtagaaat	gccgacggcg	cgccgagcac	tcgtgcttgt	gtttatacaa	7980
gcgtccgag	tgctcgcaac	gctgcacggg	atgcacgtgc	tgacagagct	gtacctgggt	8040
tcctttgacg	aggaatttca	gtgggcagtg	gagcgtggtc	ggctgcatct	ggtgctgtac	8100
tacgtcctgg	ccatcggcgt	ggccatcgtc	tgcttcgatg	gtggtcatgc	tgacgaggcc	8160
gcgcgggagg	caggtccaga	cctcggctcg	gacgggtcgg	agagcgagga	cgagggcgcg	8220
caggccggag	ctgtccaggg	tcctgagacg	ctgcggagtc	aggtcagtg	gcagcggcgg	8280
cgcgcggttg	acttgagga	gctttttccag	ggcgcgcggg	aggtccagat	ggtacttgat	8340
ctccacggcg	ccgttggtgg	cgacgtccac	ggcttgacgg	gtcccgtgcc	cctggggcg	8400
caccaccgtg	ccccgtttct	tcttggggtg	tgccggcgcc	ggctccatgc	ttagaagcgg	8460
cggcgaggac	gcgcgccggg	cggcaggggc	ggctcggggc	ccggaggcag	gggcggcagg	8520
ggcacgtcgg	cgccgcgcgc	gggcagggtc	tggtactgcg	cccggagaag	actggcgtga	8580
gcgacgacgc	gacggttgac	gtcctggatc	tgacgcctct	gggtgaaggc	cacgggaccc	8640
gtgagtttga	acctgaaaga	gagttcgaca	gaatcaatct	cggtatcggt	gacggcgggc	8700
tgccgcagga	tctcttgac	gtcgcggcag	ttgtcctggg	agggcatctc	ggtcatgaac	8760
tgctcgatct	cctcctcctg	aagggtctcc	cgaccggcgc	gctcgacggg	ggccgcgagg	8820
tcgttgagg	tgccggccat	gagctgcgag	aaggcggtca	tgccggcctc	gttcagacg	8880
cggctgtaga	ccacgggtcc	gtcgggggtc	cgcgcgcgca	tgaccacctg	ggcgagggtg	8940
agctcgacgt	ggcgcggtga	gaccgcgtag	ttgcagaggc	gctggtagag	gtagttgagc	9000
gtggtggcga	tggtgctcgg	gacgaagaag	tacatgatcc	agcggcgagg	cggcatctcg	9060
ctgacgtcgc	ccaggggttc	caagcgctcc	atggcctcgt	agaagtccac	ggcgaagttg	9120
aaaaactggg	agttgcgcgc	cgagacggtc	aactcctcct	ccagaagacg	gatgagctcg	9180
gcgatgggtg	cgcgcacctc	gcgctcgaag	gccccggggg	gctcctcttc	ttccatctcc	9240
tcctcctctt	ccatctcctc	cactaacatc	tcttctactt	cctcctcagg	aggcgggcgg	9300
gggggagggg	ccctgcgtcg	ccggcgggcg	acgggcagac	ggtcgatgaa	gcgctcgatg	9360
gtctccccgc	gccggcgacg	catgggtctg	gtgacggcgc	gcccgtcctc	gcggggccgc	9420
agcgtgaaga	cgccgcccg	catctccagg	tgcccgccgg	gggggtctcc	gttgggcagg	9480

gagagggcgc tgacgatgca tcttatcaat tggcccgtag ggactccgcg caaggacctg 9540
agcgtctcga gatccacggg atccgaaaac cgctgaacga aggccttcgag ccagtcgcag 9600
tcgcaaggta ggctgagccc ggtttcttgt tcttcgggta tttggtcggg aggcgggcgg 9660
gcatgctgc tggatgatgaa gttgaagtag gcggtcctga gacggcggat ggtggcgagg 9720
agcaccaggt ccttgggccc ggcttgctgg atgcgcagac ggtcggccat gccccaggcg 9780
tggctctgac acctggcgag gtcctttag tagtcctgca tgagccgctc cacgggcacc 9840
tcctcctcgc ccgcgcggcc gtgcatgcgc gtgagcccga acccgcgctg cggctggacg 9900
agcgccaggt cggcgacgac gcgctcggcg aggatggcct gctggatctg ggtgagggtg 9960
gtctggaagt cgtcgaagtc gacgaagcgg tggtaggctc cgggtgttgat ggtgtaggag 10020
cagttggcca tgacggacca gttgacggtc tggtagggcg ggcgcacgag ctcgtggtac 10080
ttgaggcgcg agtaggcgcg cgtgtcgaag atgtagtcgt tgcagggtgcg cacgaggtag 10140
tggatatcca cgaggaagtg cggcggcggc tggcggtaga gcggccatcg ctcggtggcg 10200
ggggcgccgg gcgcgaggtc ctcgagcatg aggcgggtgg agccgtagat gtacctggac 10260
atccaggtga tgccggcggc ggtgggtggag gcgcgcggga actcgcgac gcggttccag 10320
atgttgcgca gcggcaggaa gtagttcatg gtggccgcgg tctggcccgt gaggcgcgcg 10380
cagtcgtgga tgctctagac atacgggcaa aaacgaaagc ggtcagcggc tcgactccgt 10440
ggcctggagg ctaagcgaac gggttgggct gcgcgtgtac cccggttcga gtccctgctc 10500
gaatcaggct ggagccgcag ctaacgtggt actggcactc ccgtctcgac ccaagcctgc 10560
taacgaaacc tccaggatac ggaggcgggt cgttttgggc attttcgtca ggccggaaat 10620
gaaactagta agcgcggaac gcggccgtcc gcgatggctc gctgccgtag tctggagaaa 10680
gaatcgccag ggttgcggtg cgggtgtgcc cggttcgagc ctcagcgctc ggcgccggcc 10740
ggattccgcg gctaacgtgg gcgtggctgc cccgtcgttt ccaagacccc ttagccagcc 10800
gacttctcca gttacggagc gagcccctct ttttcttgtg tttttgccag atgcatcccg 10860
tactgcggca gatgcgcccc caccctccac cacaaccgcc cctaccgcag cagcagcaac 10920
agccggcgct tctgcccccg ccccgagcag agcagccagc cactaccgcg gcggccgccg 10980
tgagcggagc cggcgttcag tatgacctgg ccttggaaga gggcgagggg ctggcgcggc 11040
tgggggcgct gtcgccggag cggcacccgc gcgtgcagat gaaaaggag gctcgcgagg 11100
cctacgtgcc caagcagaac ctgttcagag acaggagcgg cgaggagccc gaggagatgc 11160
gcgcctcccg cttccacgcg gggcgggagc tgccggcgcg cctggaccga aagcgggtgc 11220
tgagggacga ggatttcgag gcggacgagc tgacggggat cagccccgcg cgcgcgcacg 11280
tggccgcggc caacctgggtc acggcgtagc agcagaccgt gaaggaggag agcaacttcc 11340
aaaaatcctt caacaaccac gtgcgcacgc tgatgcgcgc cgaggaggtg accctgggccc 11400
tgatgcacct gtgggacctg ctggaggcca tcgtgcagaa cccacgagc aagccgctga 11460
cggcgagct gtttctgggtg gtgcagcaca gtcgggacaa cgagacgttc agggaggcgc 11520

tgctgaatat caccgagccc gagggccgct ggctcctgga cctggtgaac attctgcaga 11580
gcatcgtggt gcaggagcgc gggctgccgc tgtccgagaa gctggcggcc atcaacttct 11640
cgggtgctgag cctgggcaag tactacgcta ggaagatcta caagaccccg tacgtgcca 11700
tagacaagga ggtgaagatc gacgggtttt acatgcgcat gaccctgaaa gtgctgaccc 11760
tgagcgacga tctgggggtg taccgcaacg acaggatgca ccgcgcggtg agcgcagacc 11820
gccggcgcgga gctgagcgac caggagctga tgcacagcct gcagcgggccc ctgaccgggg 11880
ccgggaccga gggggagagc tactttgaca tgggcgcgga cctgcgctgg cagcctagcc 11940
gccgggcctt ggaagctgcc ggcggttccc cctacgtgga ggaggtggac gatgaggagg 12000
aggagggcga gtacctggaa gactgatggc gcgaccgtat ttttgctaga tgcagcaaca 12060
gccaccgccc cctcctgatc ccgcgatgcg ggcggcgctg cagagccagc cgtccggcat 12120
taactcctcg gacgattgga cccaggccat gcaacgcata atggcgctga cgaccgcaa 12180
tcccgaagcc tttagacagc agcctcaggc caaccgactc tcggccatcc tggaggccgt 12240
ggtgccctcg cgctgaacc ccacgcacga gaaggtgctg gccatcgtga acgcgctggt 12300
ggagaacaag gccatccgcg gcgacgaggc cgggctggtg tacaacgcgc tgctggagcg 12360
cgtggcccg c tacaacagca ccaacgtgca gacgaacctg gaccgcatgg tgaccgacgt 12420
gcgcgaggcg gtgtcgcagc gcgagcggtt ccaccgcgag tcgaacctgg gctccatggt 12480
ggcgctgaac gccttctga gcacgcagcc cgccaacgtg ccccggggccc aggaggacta 12540
caccaacttc atcagcgcg cgcggtgat ggtggccgag gtgccccaga gcgaggtgta 12600
ccagtcgggg ccggactact tcttcagac cagtcgccag ggcttgaga ccgtgaacct 12660
gagccaggct ttcaagaact tgcaaggact gtggggcggtg caggccccgg tcggggaccg 12720
cgcgacggtg tcgagcctgc tgacgccgaa ctgcgcctg ctgctgctgc tgggtggcgcc 12780
cttcacggac agcggcagcg tgagccgcga ctcgctacctg ggctacctgc ttaacctgta 12840
ccgcgaggcc atcgggcagg gcgacgtgga cgagcagacc taccaggaga tcaccacgt 12900
gagccgcgcg ctgggcccagg aggaccggg caacctggag gccaccctga acttctgct 12960
gaccaaccgg tcgcagaaga tcccgcacca gtacgcgctg agcaccgagg aggagcgcat 13020
cctgcgctac gtgcagcaga gcgtggggct gttcctgatg caggaggggg ccacgcccag 13080
cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc atgtacgccc gcaaccgccc 13140
gttcatcaat aagctgatgg actacttgca tcgggcggcc gccatgaact cggactactt 13200
taccaacgcc atcttgaacc cgcactggct cccgccgccc gggttctaca cgggcgagta 13260
cgacatgccc gacccaacg acgggttctt gtgggacgac gtggacagca gcgtgttctc 13320
gccgcgcccc accaccacca ccgtgtggaa gaaagagggc ggggaccggc ggccgtctc 13380
ggcgctgtcc ggtcgcgcgg gtgctgcgcg ggcggtgccc gaggccgcca gccccttccc 13440
gagcctgccc ttttcgctga acagcgtgcg cagcagcgag ctgggtcggc tgacgcggcc 13500
gcgcctgctg ggcgaggagg agtacctgaa cgactccttg cttcggcccc agcgcgagaa 13560
gaacttcccc aataacggga tagagagcct ggtggacaag atgagccgct ggaagacgta 13620

cgcgacagag cacagggacg agccccgagc tagcagcagc accggcgcca cccgtagacg 13680
 ccagcggcac gacaggcagc ggggtctggt gtgggacgat gaggattccg ccgacgacag 13740
 cagcgtgttg gacttggggtg ggagtggtgg tggttaacccg ttcgctcacc tgcgcccccg 13800
 tatcgggagc ctgatgtaag aatctgaaaa aataaaaagac ggtactcacc aaggccatgg 13860
 cgaccagcgt gcgttcttct ctgttggttg tagtagtatg atgaggcgcg tgtaccggga 13920
 gggctctcct ccctcgtacg agagcgtgat gcagcaggcg gtggcgggcg cgatgcagcc 13980
 cccgctggag gcgccttacg tgcccccgcg gtacctggcg cctacggagg ggcggaacag 14040
 cattcgttac tcggagctgg cacccttgta cgataccacc cggttgtacc tgggtggacaa 14100
 caagtcggcg gacatcgctt cgctgaacta ccagaacgac cacagcaact tcctgaccac 14160
 cgtggtgcag aacaacgatt tcacccccac ggaggccagc acccagacca tcaactttga 14220
 cgagcgctcg cgggtggggcg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt 14280
 gaacgagttc atgtacagca acaagttcaa ggcgcgggtg atggtctcgc gcaagacccc 14340
 caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt ggggtggagtt 14400
 tgagctgccc gagggcaact tctcggtgac catgaccatc gatctgatga acaacgccat 14460
 catcgacaac tacttggcgg tggggcgcca gaacggggtg ctggagagcg acatcggcgt 14520
 gaagttcgac acgcgcaact tccggctggg ctgggacccc gtgaccgagc tgggtgatgcc 14580
 gggcgtgtac accaacgagg ccttccaccc cgacatcgtc ctgctgcccc gctgcggcgt 14640
 ggacttcacc gagagccgcc tcagcaacct gctgggcatc cgcaagcggc agcccttcca 14700
 ggagggcttc cagatcctgt acgaggacct ggaggggggc aacatccccg cgctgctgga 14760
 cgtggacgcc tacgagaaaa gcaaggagga tagcgccgcc gcggcgaccg cagccgtggc 14820
 caccgcctct accgaggtgc ggggcgataa ttttgtagc gccgagacac tggcagcggc 14880
 cgaggcggct gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagcaagga 14940
 gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtacctggc 15000
 ctacaactac ggcgaccccg agaaggcggt gcgctcctgg acgctgctca ccacctcgga 15060
 cgtcacctgc ggcgtggagc aagtctactg gtcgctgccc gacatgatgc aagaccgggt 15120
 caccttccgc tccacgcgtc aagttagcaa ctacccggtg gtgggcgccg agctcctgcc 15180
 cgtctactcc aagagcttct tcaacgagca ggccgtctac tcgcagcagc tgcgcgcctt 15240
 cacctcgtc acgcacgtct tcaaccgctt ccccgagaac cagatcctcg ttcgccccgc 15300
 cgccccacc attaccaccg tcagtgaaaa cgcttctgct ctacagatc acgggaccct 15360
 gccgctgctc agcagtatcc ggggagtgca gcgcgtgacc gtcactgacg ccagacgccg 15420
 cacctgcccc tacgtctaca agggcctggg cgtagtcgcg ccgcgcgtcc tctcgagccg 15480
 caccttctaa aaaatgtcca ttctcatctc gccagtaat aacaccggtt ggggcctgcg 15540
 cgccccagc aagatgtacg gaggcgctcg ccaacgctcc acgcaacacc ccgtgcgcgt 15600
 gcgcgggcac ttccgcgctc cctggggcgcg cctcaaggcg cgcgctgcgt cgcgaccac 15660

cgtcgacgac gtgatcgacc aggtggtggc cgacgcgcgc aactacacgc ccgccgccgc 15720
 gcccgctctcc accgtggacg ccgtcatcga cagcgtggtg gccgacgcgc gccggtacgc 15780
 ccgcgccaag agccggcggc ggcgcacgc ccggcggcac cggagcacc ccgccatgcg 15840
 cgcggcgcga gccttgctgc gcagggccag gcgcacggga cgcagggcca tgctcagggc 15900
 ggccagacgc gcggcctccg gcagcagcag cgccggcagg acccgagac gcgcggccac 15960
 ggccggcggc gcggccatcg ccagcatgtc ccgcccgcg cgcggaacg tgtactgggt 16020
 gcgcgacgcc gccaccggtg tgcgcgtgcc cgtgcgcacc cgccccctc gcacttgaag 16080
 atgctgactt cgcgatgttg atgtgtccca gcggcgagga ggatgtccaa gcgcaaattc 16140
 aaggaagaga tgctccaggt catcgcgcct gagatctacg gcccggcggc ggtgaaggag 16200
 gaaagaaagc cccgcaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgtg 16260
 gacggactgg tggagtttgt gcgcgagttc gccccccggc ggcgcgtgca gtggcgcggg 16320
 cggaaagtga aaccggtgct gcgacccggc accacggtgg tcttcacgcc cggcgagcgt 16380
 tccggtccg cctccaagcg ctctacgac gaggtgtacg gggacgagga catcctcgag 16440
 caggcggccg aacgtctggg cgagtttgct tacggcaagc gcagccgcc cgcgcccttg 16500
 aaagaggagg cgggtgtccat cccgctggac cacggcaacc ccacgccgag cctgaagccg 16560
 gtgaccctgc agcaggtgct gcctggtgcg gcgccgcgcc ggggcttcaa gcgcgagggc 16620
 ggcgaggatc tgtaccgcac catgcagctg atggtgcccc agcgccagaa gctggaggac 16680
 gtgctggagc acatgaagggt ggaccccgag gtgcagcccg aggtcaagggt gcggcccatc 16740
 aagcaggtgg ccccgggcct gggcgtgcag accgtggaca tcaagatccc cacggagccc 16800
 atggaaacgc agaccgagcc cgtgaagccc agcaccagca ccatggagggt gcagacggat 16860
 ccctggatgc cggcaccggc ttccaccacc cgccgaagac gcaagtacgg cgcgcccagc 16920
 ctgctgatgc ccaactacgc gctgcacct tccatcatcc ccacgccggg ctaccgcggc 16980
 acgcgcttct accgcggcta caccagcagc cgccgccgca agaccaccac ccgccgccgc 17040
 cgtcgtcgca cccgccgcag cagcaccgcg acttccgccg ccgccctgggt gcggagagtg 17100
 taccgcagcg ggcgcgagcc tctgaccctg ccgcgcgcgc gctaccacc gagcatcgcc 17160
 atttaactac cgctcctac ttgcagatat ggccctcaca tgccgcctcc gcgtcccat 17220
 tacgggctac cgaggaagaa agccgcgccg tagaaggctg acggggaacg ggctgcgtcg 17280
 ccatcaccac cggcggcggc gcgccatcag caagcgggtg gggggagggt tcctgcccgc 17340
 gctgatgccc atcatcgccg cggcgatcgg ggcgatcccc ggcatagctt ccgtggcgg 17400
 gcaggcctct cagcgccact gagacacagc ttggaaaatt tgtaataaaa aatggactga 17460
 cgctcctgggt cctgtgatgt gtgttttttag atggaagaca tcaatttttc gtccctggca 17520
 ccgcgacacg gcacgcggcc gtttatgggc acctggagcg acatcggcaa cagccaactg 17580
 aacggggggc cttcaattg gagcagtctc tggagcgggc ttaagaattt cgggtccacg 17640
 ctcaaaacct atggcaacaa ggcgtggaac agcagcacag ggcaaggcgt gagggaag 17700
 ctgaaagagc agaacttcca gcagaagggt gtcgatggcc tggcctcggg catcaacggg 17760

gtggtggacc tggccaacca ggccgtgcag aaacagatca acagccgcct ggacgcggtc 17820
 ccgcccgcgg ggtccgtgga gatgccccag gtggaggagg agctgcctcc cctggacaag 17880
 cgcggcgaca agcgaccgcg tcccgcgcg gaggagacgc tgctgacgca cacggacgag 17940
 ccgccccctg acgaggaggc ggtgaaactg ggtctgcca ccacgcggcc cgtggcgctc 18000
 ctggccaccg ggggtgctgaa acccagcagc agcagcagcc agcccgcgac cctggacttg 18060
 cctccgcctg cttcccgccc ctccacagtg gctaagcccc tgccgccggt ggccgtcgcg 18120
 tcgcgcgccc cccgaggccg ccccaggcg aactggcaga gcactctgaa cagcatctgt 18180
 ggtctgggag tgcagagtgt gaagcgccgc cgctgctatt aaaagacact gtagcgctta 18240
 acttgcttgt ctgtgtgtat atgtatgtcc gccgaccaga aggaggagga agaggcgct 18300
 cgccgagttg caagatggcc accccatcga tgctgccccca gtgggcgtac atgcacatcg 18360
 ccggacagga cgcttcggag tacctgagtc cgggtctggt gcagttcgcc cgcgccacag 18420
 acacctactt cagtctgggg aacaagttta ggaacccac ggtggcgccc acgcacgatg 18480
 tgaccaccga ccgcagccag cggctgacgc tgcgcttcgt gcccgaggac cgcgaggaca 18540
 acacctactc gtacaaagtg cgctacacgc tggccgtggg cgacaaccgc gtgctggaca 18600
 tggccagcac ctactttgac atccgcggcg tgctggatcg gggccctagc ttcaaaccct 18660
 actccggcac cgcttacaac agcctggctc ccaagggagc gcccaacact tgccagtgga 18720
 catataaagc tgatggtgat actggtacag aaaaaaccta tacatatgga aatgcgcctg 18780
 tgcaaggcat tagtattaca aaagatggta ttcaacttgg aactgacact gatgatcagc 18840
 ccatttatgc agataaaact tatcaaccag agcctcaagt gggatgatgct gaatggcatg 18900
 acatcactgg tactgatgaa aaatatggag gcagagctct caagcctgac accaaaatga 18960
 agccctgcta tggttctttt gccaaagccta ccaataaaga aggaggctcag gcaaatgtga 19020
 aaaccgaaac aggcggtacc aaagaatatg acattgacat ggcattcttc gataatcgaa 19080
 gtgcagctgc ggctggcctg gccccagaaa ttgttttgta tactgagaat gtggatctgg 19140
 aaactccaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca 19200
 atttgggtca gcagtccatg cccaacagac ccaactacat tggctttaga gacaacttta 19260
 tcgggctcat gtactacaac agcactggca acatgggcgt gctggctggg caggcctccc 19320
 agctgaatgc tgtggtggac ttgcaggaca gaaacactga actgtcctac cagctcttgc 19380
 ttgactctct gggcgacaga accaggtatt tcagtatgtg gaatcaggcg gtggacagct 19440
 atgaccccga tgtgcgcat attgaaaatc acggtgtgga ggatgaactc cctaactatt 19500
 gcttccccct ggatgctgtg ggtagaactg atacttacca gggaattaag gccaatggtg 19560
 ctgatcaaac cacctggacc aaagatgata ctgttaatga tgctaatgaa ttgggcaagg 19620
 gcaatccttt cgccatggag atcaacatcc aggcgaacct gtggcggaac ttcctctacg 19680
 cgaacgtggc gctgtacctg cccgactcct acaagtacac gccggccaac atcacgctgc 19740
 cgaccaacac caacacctac gattacatga acggccgcgt ggtggcgccc tcgctgggtg 19800

acgcctacat caacatcggg gcgcgctggt cgctggaccc catggacaac gtcaacccct 19860
tcaaccacca ccgcaacgcg ggcctgcgct accgctccat gtccttgggc aacgggcgct 19920
acgtgcccctt ccacatccag gtgccccaaa agttcttcgc catcaagagc ctcttgcctc 19980
tgcccgggtc ctacacctac gagtggaact tccgcaagga cgtcaacatg atcctgcaga 20040
gctccctcgg caacgacctg cgcacggacg gggcctccat cgccttcacc agcatcaacc 20100
tctacgccac cttcttcccc atggcgcaca acaccgcctc cacgctcgag gccatgctgc 20160
gcaacgacac caacgaccag tccttcaacg actacctctc ggcgggcaac atgctctacc 20220
ccatcccggc caacgccacc aacgtgcca tctccatccc ctgcgcgaac tgggcccgcct 20280
tccgcggatg gtccttcacg cgcctcaaga cccgcgagac gccctcgcct ggctccgggt 20340
tcgacccta cttcgtctac tcgggtcca tcccctacct cgacggcacc ttctacctca 20400
accacacctt caagaaggct tccatcacct tcgactcctc cgtcagctgg cccggcaacg 20460
accgcctcct gacgccaac gagttcgaaa tcaagcgcac cgtcgacgga gagggtaca 20520
acgtggccca gtgcaacatg accaaggact ggttcctggt ccagatgctg gccactaca 20580
acatcggtc ccagggttc tacgtgcccg agggctacaa ggaccgcatg tactccttct 20640
tccgcaactt ccagcccag agccgccagg tcgtggacga ggtcaactac aaggactacc 20700
aggccgtcac cctggcctac cagcacaaca actcgggctt cgtcggctac ctgcgcca 20760
ccatgcgcca gggacagccc taccgccca actacccta cccgctcatc ggcaagagcg 20820
ccgtcgccag cgtcaccag aaaaagttcc tctgcgaccg ggtcatgtgg cgcattccct 20880
tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgctctacg 20940
ccaactccgc ccacgcgcta gacatgaatt tcgaagtcga ccccatggat gagtccacc 21000
ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg 21060
tcatcgaggc cgtctacctg cgcacgccct tctcggccgg caacgccacc acctaagccc 21120
cgctcttgct tcttgcaaga tgacggcctg tgccgggtcc ggcgagcagg agctcagggc 21180
catcctccgc gacctgggt gcgggccctg cttcctgggc accttcgaca agcgcttccc 21240
gggattcatg gccccgcaca agctggcctg cgccatcgtc aacacggccg gccgcgagac 21300
cgggggcgag cactggctgg ccttcgcctg gaaccgcgc tcccacacct gctacctctt 21360
cgacccttc gggttctcgg acgagcgct caagcagatc taccagttcg agtacgaggg 21420
cctgctgcgc cgcagcgcgc tggccaccga ggaccgctgc gtcaccctgg aaaagtccac 21480
ccagaccgtg cagggtccgc gctcggccgc ctgcgggtc ttctgctgca tgttctgca 21540
cgccttcgtg cactggcccc accgccccat ggacaagaac cccaccatga acttgctgac 21600
gggggtgccc aacggcatgc tccagtcgcc ccagggtgaa cccaccctgc gccgcaacca 21660
ggaggcgctc taccgcttcc tcaacgcca ctccgcctac ttctgctccc accgcgcgcg 21720
catcgagaag gccaccgcct tcgaccgat gaatcaagac atgtaaaccg tgtgtgtatg 21780
tgaatgcttt attcataata aacagcacat gtttatgcca ctttttctga ggctctgact 21840
ttatttagaa atcgaagggg ttctgcccgc tctcggcggt ccccgcgggc agggatacgt 21900

tgcggaactg	gtacttgggc	agccacttga	actcggggat	cagcagcttc	ggcacgggga	21960
ggtcggggaa	cgagtcgctc	cacagcttgc	gcgtgagttg	cagggcgccc	agcaggtcgg	22020
gcgcggagat	cttgaaatcg	cagttgggac	ccgcgttctg	cgcgcgggag	ttgcggtaca	22080
cggggtttga	gcaactggaac	accatcaggg	ccgggtgctt	cacgctcgcc	agcaccgtcg	22140
cgtcggtgat	gccctccacg	tccagatcct	cggcgttggc	catcccgaag	ggggtcattct	22200
tgcaggctctg	ccgccccatg	ctgggcacgc	agccgggctt	gtggttgcaa	tcgcagtgca	22260
gggggatcag	catcatctgg	gcctgctcgg	agctcatgcc	cgggtacatg	gccttcattga	22320
aagcctccag	ctggcggaag	gcctgctcgg	ccttgccgcc	ctcgggtgaag	aagacccccgc	22380
aggacttgct	agagaactgg	ttggtggcgc	agccggcgctc	gtgcacgcag	cagcgcgctg	22440
cgttggttggc	cagctgcacc	acgctgcgcc	cccagcggtt	ctgggtgatc	ttggcccggg	22500
cggggtttctc	cttcagcgcg	cgctgcccgt	tctcgctcgc	cacatccatc	tcgatcgtgt	22560
gctccttctg	gatcatcacg	gtcccgtgca	ggcatcgcag	cttgccctcg	gcctcgggtgc	22620
accggtgcag	ccacagcgcg	cagccgggtgc	actcccagtt	cttggtgggcg	atctggggagt	22680
gcgagtgac	gaagccctgc	aggaagcggc	ccatcatcgt	ggtcagggtc	ttgttgctgg	22740
tgaaggtcag	cgggatgccg	cgggtgctct	cgttcacata	cagggtggcag	atgcggcggg	22800
acacctcgcc	ctgctcgggc	atcagctgga	aggcggactt	caggctcgctc	tccacgcggg	22860
accggtccat	cagcagcgctc	atgacttcca	tgcccttctc	ccaggccgag	acgatcggca	22920
ggctcagggg	gttcttcacc	gccgttgctc	tcttagtcgc	cgccgctgag	gtcagggggg	22980
cgttctcgctc	caggggtctca	aacactcgct	tgccgtcctt	ctcgggtgatg	cgcacggggg	23040
gaaagctgaa	gcccacggcc	gccagctcct	cctcggcctg	cctttcgtcc	tcgctgtcct	23100
ggctgatgtc	ttgcaaaggc	acatgcttgg	tcttgcgggg	tttctttttg	ggcggcagag	23160
gcggcggcgg	agacgtgctg	ggcgagcgcg	agttctcgct	caccacgact	atttcttctt	23220
cttgcccgctc	gtccgagacc	acgcggcggg	aggcatgcct	cttctggggc	agaggcggag	23280
gcgacgggct	ctcgcggttc	ggcgggcggc	tggcagagcc	ccttcgcgct	tcgggggtgc	23340
gctcctggcg	gcgctgctct	gactgacttc	ctccgcggcc	ggccattgtg	ttctcctagg	23400
gagcaacaag	catggagact	cagccatcgt	cgccaacatc	gccatctgcc	cccgccgccg	23460
ccgacgagaa	ccagcagcag	aatgaaagct	taaccgcccc	gccgcccagc	cccacctccg	23520
acgccgccgc	ggccccagac	atgcaagaga	tggaggaatc	catcgagatt	gacctgggct	23580
acgtgacgcc	cgcgagacac	gaggaggagc	tggcagcgcg	cttttcagcc	ccggaagaga	23640
accaccaaga	gcagccagag	caggaagcag	agagcgagca	gcagcaggct	gggctcagac	23700
atggcgacta	cctgagcggg	gcagaggacg	tgctcatcaa	gcatctggcc	cgccaatgca	23760
tcacgtgcaa	ggacgcgctg	ctcgaccgcg	ccgaggtgcc	cctcagcgctg	gcggagctca	23820
gccgcgccta	cgagcgcaac	ctcttctcgc	cgcgcggtgcc	ccccaaagcg	cagcccaacg	23880
gcacctgcga	gccaacccg	cgctcaact	tctaccgggt	cttcgcgggtg	cccagggccc	23940

tggccaccta ccacctcttt ttcaagaacc aaaggatccc cgtctcctgc cgcgccaacc 24000
gcacccgcgc cgacgccctg ctcaacctgg gtcccggcgc ccgcctacct gatatcgctt 24060
ccttggaaga ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggcccgcga 24120
acgctctgca aggaagcgga gaggagcatg agcaccacag cgccctggtg gagttggaag 24180
gcgacaacgc gcgcctggcg gtgctcaagc gcacggtcga gctgaccacac ttcgcctacc 24240
cggcgctcaa cctgcccccc aaggctcatga gcgccgtcat ggaccaggtg ctcatcaagc 24300
gcgcctcgcc cctctcggat gaggacatgc aggaccccga gagctcggac gagggcaagc 24360
ccgtggtcag cgacgagcag ctggcgcgct ggctgggagc gagtagcacc ccccagagct 24420
tggaagagcg gcgcaagctc atgatggccg tggctcctggt gaccgtggag ctggagtgtc 24480
tgcgccgctt cttcgccgac gcagagaccc tgcgcaaggt cgaggagaac ctgcactacc 24540
tcttcaggca cgggtttgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc 24600
tggctctcta catgggcatc ctgcacgaga accgcctggg gcagaacgtg ctgcacacca 24660
ccctgcgcgg ggaggcccg gcgcactaca tccgcgactg cgtctacctg tacctctgcc 24720
acacctggca gacgggcatg ggcgtgtggc agcagtgcct ggaggagcag aacctgaaag 24780
agctctgcaa gctcctgcag aagaacctga aggcctgtg gaccgggttc gacgagcgca 24840
ccaccgcctc ggacctggcc gacctcatct tccccgagcg cctgcggctg acgctgcgca 24900
acggactgcc cgactttatg agtcaaagca tgttgcaaaa ctttcgctct ttcatacctg 24960
aacgctccgg gatcctgccc gccacctgct ccgcgctgcc ctcggaactc gtgccgctga 25020
ccttccgcga gtgccccccg ccgctctgga gccactgcta cctgctgcgc ctggccaact 25080
acctggccta ccactcggac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc 25140
actgccgctg caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga 25200
gcgagacca gatcatcggc accttcgagt tgcaaggccc cggcgagggc aaggggggtc 25260
tgaaactcac cccggggctg tggacctcg cctacttgcg caagttcgtg cccgaggact 25320
accatccctt cgagatcagg ttctacgagg accaatcca gccgcccaag gccgaactgt 25380
cggcctgcgt catcaccag ggggccatcc tggccaatt gcaagccatc cagaaatccc 25440
gccaagaatt tctgctgaaa aagggccacg ggggtctacct ggacccccag accggagagg 25500
agctcaaccc cagcttcccc caggatgccc cgaggaaagca gcaagaagct gaaagtggag 25560
ctgccgccgc cggaggattt ggaggaagac tgggagagca gtcaggcaga ggaggaggag 25620
atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctggaagac 25680
gaggtggagg aggaggcaga ggaagaagca gccgccgcca gaccgtcgtc ctcggcggag 25740
aaagcaagca gcacggatac catctccgct ccgggtcggg gtcgcggcga ccgggcccac 25800
agtaggtggg acgagaccgg gcgcttcccc aacccacca cccagaccgg taagaaggag 25860
cggcagggat acaagtcctg gcgggggcac aaaaacgcca tcgtctcctg cttgcaagcc 25920
tgcgggggca acatctcctt caccgcgcgc tacctgtctt tccaccgagg ggtgaacttc 25980
ccccgcaaca tcttgatta ctaccgtcac ctccacagcc cctactactg tttccaagaa 26040

gaggcagaaa	cccagcagca	gcagaaaacc	agcggcagca	gcagctagaa	aatccacagc	26100
ggcggcaggt	ggactgagga	tcgcagcgaa	cgagccggcg	cagacccggg	agctgaggaa	26160
ccggatcttt	cccaccctct	atgccatctt	ccagcagagt	cgggggcagg	agcaggaact	26220
gaaagtcaag	aaccgttctc	tgcgctcgct	caccgcgagt	tgtctgtatc	acaagagcga	26280
agaccaactt	cagcgcactc	tcgaggacgc	cgaggctctc	ttcaacaagt	actgcgcgct	26340
cactcttaaa	gagtagcccc	cgccccccca	cacacggaaa	aaggcgggaa	ttacgtcacc	26400
acctgcgccc	ttcggccgac	catcatcatg	agcaaagaga	ttcccacgcc	ttacatgtgg	26460
agctaccagc	cccagatggg	cctggccgcc	ggcggccgcc	aggactactc	caccgcgatg	26520
aactggctca	gcgccggggc	cgcgatgata	tcacgggtga	atgacatccg	cgcccggcga	26580
aaccagatac	tcctagaaca	gtcagcgata	accgccacgc	cccgccatca	ccttaatccg	26640
cgtaattggc	ccgccgccct	ggtgtaccag	gaaattcccc	agcccacgac	cgtactactt	26700
ccgcgagacg	cccaggccga	agtccagctg	actaactcag	gtgtccagct	ggccggcggc	26760
gccgccctgt	gtcgtcaccg	ccccgctcag	ggtataaagc	ggctgggtgat	ccgaggcaga	26820
ggcacacagc	tcaacgacga	ggtggtgagc	tcttcgctgg	gtctgcgacc	tgacggagtc	26880
ttccaactcg	ccggatcggg	gagatcttcc	ttcacgcctc	gtcaggccgt	cctgactttg	26940
gagagttcgt	cctcgagcc	ccgctcgggt	ggcatcggca	ctctccagtt	cgtggaggag	27000
ttcactccct	cggctactt	caacccttc	tccggctccc	ccggccacta	cccggacgag	27060
ttcatcccga	acttcgacgc	catcagcgag	tcggtggacg	gctacgattg	aatgtcccat	27120
ggtggcgag	ctgacctagc	tcggcttcga	cacctggacc	actgccgccg	cttccgctgc	27180
ttcgctcggg	atctcgccga	gtttgcctac	tttgagctgc	ccgaggagca	ccctcagggc	27240
ccggccccacg	gagtgcggat	catcgtcgaa	gggggcctcg	actcccacct	gcttcggatc	27300
ttcagccagc	gaccgatcct	ggtcgagcgc	gagcaaggac	agacccttct	gaccctgtac	27360
tgcattctgca	accaccccg	cctgcatgaa	agtctttggt	gtctgctgtg	tactgagtat	27420
aataaaaagct	gagatcagcg	actactccgg	actcgattgt	ggtgttcctg	ctatcaaccg	27480
gtccctgttc	ttcaccggga	acgagaccga	gctccagctt	cagtgtgtaagc	cccacaagaa	27540
gtacctcacc	tggtgttcc	agggtcctcc	gatcgccggt	gtcaaccact	gcgacaacga	27600
cggagtcctg	ctgagcggcc	ccgccaaact	tactttttcc	acccgcagaa	gcaagctcca	27660
gctctttcaa	cccttcctcc	ccgggacctc	tcagtgcgtc	tcgggaccct	gccatcacac	27720
cttccacctg	atcccgaata	ccacagcgcc	gctccccgct	actaacaacc	aaactaccca	27780
ccatcgccac	cgtcgcgacc	tttctgaatc	taacactacc	acccacaccg	gaggtgagct	27840
ccgaggtcga	ccaacctctg	ggatttacta	cggccccctg	gaggtggtgg	ggttaatagc	27900
gctaggccta	gttggtgggtg	ggctttttggc	tctctgctac	ctatacctcc	cttgctgttc	27960
gtacttagtg	gtgctgtgtt	gctggtttta	gaaatggggg	agatcacctc	agtgagctgc	28020
ggtgctgctg	tggcggtggt	ggtgttttcg	attgtgggac	tgggcggcgc	ggctgtagtg	28080

aaggagaagg ccgatccctg cttgcatttc aatccccgaca attgccagct gagttttcag 28140
cccgatggca atcgggtgcgc ggtgctgac aagtgcggat gggaatgcga gaacgtgaga 28200
atcgagtaca ataacaagac tcggaacaat actctcgcgt ccgtgtggca gcccggggac 28260
cccgagtggg acaccgtctc tgtccccggg gctgacggct ccccgcgac cgtgaacaat 28320
actttcattt ttgcgcacat gtgcgacacg gtcattgtga tgagcaagca gtacgatatg 28380
tggcccccca cgaaggagaa catcgtgggc ttctccatcg cttacagcgc gtgcacggcg 28440
ctaatacccg ctatcgtgtg cctgagcatt cacatgctca tcgctattcg cccagaaaat 28500
aatgccgaaa aagagaaaaca gccataacac gttttttcac acaccttttt cagaccatgg 28560
cctctgttaa atttttgctt ttatttgcca gtctcattac tgttataagt aatgagaaac 28620
tcactattta cattggcact aaccacactt tagacggaat tccaaaatcc tcatggtatt 28680
gctattttga tcaagatcca gacttaacta tagaactgtg tggtaacaag ggaaaaaata 28740
caagcattca ttttaattaac ttttaattgcg gagacaattt gaaattaatt aatatcacta 28800
aagagtatgg aggtatgtat tactatgttg cagaaaataa caacatgcag ttttatgaag 28860
ttactgtaac taatcccacc acacctagaa caacaacaac caccaccaca aaaactacac 28920
ctgttaccac tatgcagctc actaccaata acatttttgc catgcgtcaa atggtcaaca 28980
atagcactca acccaccaca ccagtgagg aaattcccaa atccatgatt ggcattattg 29040
ttgctgtagt ggtgtgcatg ttgatcatcg ccttgtgcat ggtgtactat gccttctgct 29100
acagaaagca cagactgaac gacaagctgg aacacttact aagtgttgaa ttttaatttt 29160
ttagaaccat gaagatccta ggccttttaa ttttttctat cattacctct gctctatgca 29220
attctgacaa tgaggacgtt actgtcgttg tcggaaccaa ttatacactg aaagggtccag 29280
cgaaggggtat gctttcgtgg tattgctggg ttggaactga cgagcaacag acagagctct 29340
gcaatgctca aaaaggcaaa acctcaaatt ctaaaatctc taattatcaa tgcaatggca 29400
ctgacttagt actgctcaat gtcacgaaag catatgctgg cagctacacc tgccctggag 29460
atgatactga gaacatgatt ttttacaag tggaagtggg tgatcccact actccacctc 29520
caccaccac aactactcac accacacaca cagaacaaac cacagcagag gaggcagcaa 29580
agttagcctt gcagggtcaa gacagttcat ttgttggcat taccctaca cctgatcagc 29640
gggtgtccggg gctgctcgtc agcggcattg tcggtgtgct ttcgggatta gcagtcataa 29700
tcatctgcat gttcattttt gcttgcgtgct atagaaggct ttaccgacaa aaatcagacc 29760
cactgctgaa cctctatggt taattttttc cagagccatg aaggcagtta gcactctagt 29820
tttttgttct ttgattggca ctgttttttag tgtttagcttt ttgaaacaaa tcaatgttac 29880
tgaggggggaa aatgtgacac tggtaggcgt agagggtgct caaaatacca cctggacaaa 29940
attccatcta gatgggtgga aagaaatttg cacctggaat gtcagtactt atacatgtga 30000
aggagttaat cttaccattg tcaatgtcag ccaaattcaa aagggttgga ttaaagggca 30060
atctgttagt gttagcaata gtgggtacta taccagcat actcttatct atgacattat 30120
agttatacca ctgcctacac ctagcccacc tagcactacc acacagacaa cccacactac 30180

acaaacaacc acatacagta catcaaatca gcctaccacc actacaacag cagaggttgc 30240
 cagctcgtct ggggtccgag tggcattttt gatgttggcc ccatctagca gtcccactgc 30300
 tagtaccaat gagcagacta ctgaattttt gtccactgtc gagagccaca ccacagctac 30360
 ctcgagtgcc ttctctagca ccgccaatct atcctcgctt tcctctacac caatcagtcc 30420
 cgctactact cctacccccg ctattctccc cactcccctg aagcaaacag acggcgacat 30480
 gcaatggcag atcacctgc tcattgtgat cgggttggc atcctggccg tgttgctcta 30540
 ctacatcttc tgccgccgca ttcccaacgc gcaccgcaag ccggcctaca agcccatcgt 30600
 tgtcgggcag ccggagccgc ttcaggtgga aggggttcta aggaatcttc tcttctcttt 30660
 tacagtatgg tgattgaatt atgattccta gacaaatctt gatcactatt cttatctgcc 30720
 tcctccaagt ctgtgccacc ctgctctggt tggccaacgc cagtccagac tgtattgggc 30780
 ccttcgcctc ctacgtgctc ttgacctca tcacctgcat ctgctgctgt agcatagtct 30840
 gcctgcttat caccttcttc cagttcattg actggatctt tgtgcgcatc gcctacctgc 30900
 gccaccaccc ccagtaccgc gaccagcgag tggcgcggt gctcaggatc ctctgataag 30960
 catgcgggct ctgctacttc tcgctcttct gctgttagtg ctccccgct ccgtcgaccc 31020
 ccggaccccc acccagtccc ccgaggaggt ccgcaaagtc aaattccaag aaccctggaa 31080
 attcctcaaa tgctaccgcc aaaaatcaga catgcatccc agctggatca tgatcattgg 31140
 gatcgtgaac attctggcct gcacctcat ctcttttggt atttaccctt gctttgactt 31200
 tggttggaac tcgccagagg cgctctatct cccgcctgaa cctgacacac caccacagca 31260
 acctcaggca cacgcactac caccaccacc acagcctagg ccacaatata tgcccatatt 31320
 agactatgag gccgagccac agcgacccat gctccccgct attagttact tcaatctaac 31380
 cggcggagat gactgaccca ctggccaaca acaacgtcaa cgaccttctc ctggacatgg 31440
 acggccgcgc ctcgagcag cgactcgccc aacttcgcat tcgccagcag caggagagag 31500
 ccgtcaagga gctgcaggac ggcatagcca tccaccagt caagaaaggc atcttctgcc 31560
 tggtgaaaca ggccaagatc tcctacgagg tcaccagac cgaccatcgc ctctctacg 31620
 agtcctgca gcagcgccag aagttcacct gcctggctcg agtcaacccc atcgtcatca 31680
 cccagcagtc gggcgatacc aaggggtgca tccactgctc ctgctgactc cccgactgcg 31740
 tccacactct gatcaagacc ctctgcggcc tccgcgacct cctccccatg aactaatcac 31800
 ccccttatcc agtgaaataa agatcatatt gatgatttga gtttaataaa aataaagaat 31860
 cacttacttg aaatctgata ccaggtctct gtccatgttt tctgccaaca ccacttact 31920
 cccctcttcc cagctctggt actgcaggcc ccggcgggct gcaaacttcc tccacaccct 31980
 gaaggggatg tcaaattcct cctgtccctc aatcttcatt ttatcttcta tcagatgtcc 32040
 aaaaagcgcg tccgggtgga tgatgacttc gaccccgctt acccctacga tgcagacaac 32100
 gcaccgaccg tgcccttcat caaccccccc ttcgtctctt cagatggatt ccaagagaag 32160
 cccctggggg tgctgtccct gcgtctggcc gatcccgta ccaccaagaa cggggaaatc 32220

accctcaagc tgggagatgg ggtggacctc gactcctcgg gaaaactcat ctccaacacg 32280
 gccaccaagg ccgccgcccc tctcagtttt tccaacaaca ccatttccct taacatggat 32340
 acccctttttt acaacaacaa tggaaagtta ggcatgaaag tctactgctcc actgaagata 32400
 ctagacacag acttgctaaa aacacttggt gtagcttatg gacaagggtt aggaacaaac 32460
 accactgggtg cccttggtgc ccaactagca tccccacttg cttttgatag caatagcaaa 32520
 attgccctta atttaggcaa tggaccattg aaagtggatg caaatagact gaacatcaat 32580
 tgcaatagag gactctatgt tactaccaca aaagatgcac tgggaagcaa tataagttgg 32640
 gctaattgcta tgacatttat aggaaatgcc atgggtgtca atattgatac aaaaaaggc 32700
 ttgcaatttg gcaccactag taccgtcgca gatgttaaaa acgcttacct catacaaattc 32760
 aaacttggag ctggtctcac atttgacagc acagggtgcaa ttggtgcatg gaacaaagat 32820
 gatgacaagc ttacactatg gaccacagcc gaccctctc caaattgtca catatattct 32880
 gaaaaggatg ctaagcttac actttgcttg acaaagtgtg gcagtcagat tctgggcact 32940
 gtttccctca tagctgttga tactggcagt ttaaattcca taacagggaac agtaaccact 33000
 gctcttgtct cacttaaatt cgatgcaaatt ggagttttgc aaagcagctc aacactagac 33060
 tcagactatt ggaatttcag acaggagat gttacacctg ctgaagccta tactaatgct 33120
 ataggtttca tgcccaattc aaaagcatac cctaaaaaca caagtggagc tgcaaaaagt 33180
 cacattgttg ggaaagtgtg cctacatggg gatacaggca aaccactgga cctcattatt 33240
 actttcaatg aaacaagtga tgaattctgc acttactgta ttaactttca atggcagtg 33300
 ggggctgatc aatataaaaa tgaaacactt gccgtcagtt cattcacctt ttcctatatt 33360
 gctaaagaat aaacccact ctgtacccca tctctgtcta tggaaaaaac tctgaaacac 33420
 aaaataaaat aaagttcaag tgttttattg attcaacagt ttacaggat tcgagcagtt 33480
 atttttctc caccctcca ggacatggaa tacaccacc tctcccccg cacagccttg 33540
 aacatctgaa tgccattggt gatggacatg cttttggtct ccacgttcca cacagtttca 33600
 gagcgagcca gtctcgggtc ggtcaggag atgaaaccct ccgggcactc ccgcatctgc 33660
 acctcacagc tcaacagctg aggattgtcc tcggtggtcg ggatcacggt tatctggaag 33720
 aagcagaaga gcggcggtgg gaatcatagt ccgcgaacgg gatcggccgg tgggtgtcga 33780
 tcaggccccg cagcagtcgc tgtcgcgcc gctccgtcaa gctgctgctc aggggggtccg 33840
 ggtccaggga ctccctcagc atgatgcca cggccctcag catcagtcgt ctggtgcggc 33900
 gggcgagca gcgcatgcgg atctcgtca ggtcgtgca gtacgtgcaa cacaggacca 33960
 ccaggttggt caacagtcca tagttcaaca cgctccagcc gaaactcatc gcgggaagga 34020
 tgctaccac gtggccgtcg taccagatcc tcaggtaaatt caagtggcgc cccctccaga 34080
 acacgtgcc catgtacatg atctccttg gcatgtggcg gttcaccacc tcccgtacc 34140
 acatcacct ctggttgaa atgcagcccc ggatgatcct gcggaaccac agggccagca 34200
 ccgccccgcc cgccatgcag cgaagagacc ccgggtcccg acaatggcaa tggaggaccc 34260
 accgctcgta cccgtggatc atctgggagc tgaacaagtc tatgttggca cagcacaggc 34320

atatgctcat gcattctctc agcactctca gctcctcggg ggtcaaaacc atatcccagg 34380
 gcacggggaa ctcttgacagg acagcgaacc ccgcagaaca gggcaatcct cgcacataac 34440
 ttacattgtg catggacagg gtatcgcaat caggcagcac cgggtgatcc tccaccagag 34500
 aagcgcgggg ctcggtctcc tcacagcgtg gtaagggggc cggccgatac ggggtgatggc 34560
 gggacgcggc tgatcgtgtt cgcgaccgtg ttatgatgca gttgctttcg gacattttcg 34620
 tacttgctgt agcagaacct ggtccggggc ctgcacaccg atcgccggcg gcggtcccgg 34680
 cgcttggaac gtcgggtgtt gaagttgtaa aacagccact ctctcagacc gtgcagcaga 34740
 tctagggcct caggagtgat gaagatccca tcatgcctga tggctctaata cacatcgacc 34800
 accgtggaat gggccagacc cagccagatg atgcaatttt gttgggtttc ggtgacggcg 34860
 ggggagggaa gaacaggaag aaccatgatt aacttttaac ccaaacggtc tcggagcact 34920
 tcaaaatgaa gatcgcgagg atggcacctc tcgccccgcg tgtgttggtg gaaaataaca 34980
 gccagggtcaa aggtgatacg gttctcgaga tgttccacgg tggcttcag caaagcctcc 35040
 acgcgacat ccagaaacaa gacaatagcg aaagcgggag ggttctctaa ttcctcaatc 35100
 atcatgttac actcctgcac catccccaga taattttcat ttttccagcc ttgaatgatt 35160
 cgaactagtt cctgaggtaa atccaagcca gccatgataa agagctcgcg cagagcgccc 35220
 tccaccggca ttcttaagca caccctcata attccaagat attctgctcc tggttcacct 35280
 gcagcagatt gacaagcgga atatcaaaat ctctgccgcg atccctaagc tcctccctca 35340
 gcaataactg taagtactct ttcataatcct ctccgaaatt tttagccata ggaccaccag 35400
 gaataagatt agggcaagcc acagtacaga taaaccgaag tcctccccag tgagcattgc 35460
 caaatgcaag actgctataa gcatgctggc tagaccgggt gatattcttc agataactgg 35520
 acagaaaatc gcccaggcaa tttttaagaa aatcaacaaa agaaaaatcc tccagggtgca 35580
 cgtttagagc ctcggaaca acgatggagt aaatgcaagc ggtgcttcc agcatggtta 35640
 gttagctgat ctgtagaaaa aaacaaaaat gaacattaaa ccatgctagc ctggcgaaca 35700
 ggtgggtaaa tcgttctctc cagcaccagg caggccacgg ggtctccggc acgaccctcg 35760
 taaaaattgt cgctatgatt gaaaaccatc acagagagac gttcccgggt gccggcgatga 35820
 atgattcgac aagatgaata ccccccgga acattggcgt ccgcgagtga aaaaaagcgc 35880
 ccaaggaagc aataaggcac tacaatgctc agtctcaagt ccagcaaagc gatgccatgc 35940
 ggatgaagca caaaattctc aggtgctgac aaaatgtaat tactcccctc ctgcacaggc 36000
 agcaaagccc ccgatccctc caggtagaca tacaagcct cagcgtccat agcttaccga 36060
 gcagcagcac acaacaggcg caagagtcag agaaaggctg agctctaacc tgtccacccg 36120
 ctctctgctc aatatatagc ccagatctac actgacgtaa aggccaaagt ctaaaaatac 36180
 ccgccaata atcacacacg cccagcacac gcccagaaac cggtagacaca ctcaaaaaaa 36240
 tacgcgcact tcctcaaacg cccaaactgc cgtcatttcc gggttccac gctacgtcat 36300
 caaaattcga ctttcaaatt ccgtcgaccg ttaaaaacgt cgcccgcgcc gccctaacg 36360

gtcgccgctc ccgcagccaa tcaccgcccc gcatccccaa attcaaatac ctcatTTTgca 36420
tattaacgcg caccaaaagt ttgaggtata ttattgatga tg 36462

<210> 2
<211> 36604
<212> DNA
<213> chimpanzee adenovirus serotype Pan6

<400> 2
catcatcaat aatatacctc aaactTTTtg tgcgcgTTaa tatgcaaag agctgTTTga 60
atttggggag ggaggaaggt gattggctgc gggagcggcg accgttaggg gcggggcggg 120
tgacgTTTTg atgacgtggc tatgaggcgg agccggtTTg caagtTctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa tttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
ttccgcgtac ggtgtcaaag tccggtgTTt ttacgtaggc gtcagctgat cgccagggtg 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagtTTTct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
gtaatgTTTT cctggctact gggaacgaga ttctggaatt ggtggtggac gccatgatgg 660
gtgacgaccc tccagagccc cctaccccat ttgaggcgcc ttcgctgtac gatttgatatg 720
atctggaggt ggatgtgccc gagagcgacc ctaacgagga ggcggtgaat gatttgTTta 780
gcgatgccgc gctgctggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
cctctctcca taccgagaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg 960
aggaggcgat tcgagctgcg gtgaaccagg gagtgaaaac tgcgggcgag agcttttagcc 1020
tggaactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcataaata 1080
ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgTTt 1140
acagtaagtg tgattaactt tagttgggaa ggcagagggg gactgggtgc tgactggTTt 1200
atttatgtat atgtTTTTtt atgtgtaggt cccgtctctg acgtagatga gacccccact 1260
tcagagtgca tttcatcacc cccagaaatt ggcgaggaac cgcccgaaga tattattcat 1320
agaccagttg cagtgagagt caccgggcgg agagcagctg tggagagttt ggatgacttg 1380
ctacaggggtg gggatgaacc tttggacttg tgtaccgga aacgccccag gcactaagtg 1440
ccacacatgt gtgtttactt aaggatgatg cagtatttat aggggtgtgga gtgcaataaa 1500
atccgtgttg actttaagtg cgtgtttttat gactcagggg tggggactgt gggatatataa 1560
gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gactgtcttg 1620
gaagactttc accagactag acagttgcta gagaactcat cggagggagt ctcttacctg 1680
tggagattct gcttcggtgg gectctagct aagctagtct atagggccaa acaggattat 1740

aaggaacaat	ttgaggatat	tttgagagag	tgtcctggta	tttttgactc	tctcaacttg	1800
ggccatcagt	ctcactttta	ccagagtatt	ctgagagccc	ttgacttttc	tactcctggc	1860
agaactaccg	ccgcggtagc	cttttttgcc	tttattcttg	acaaatggag	tcaagaaacc	1920
catttcagca	gggattaccg	tctggactgc	ttagcagtag	ctttgtggag	aacatggagg	1980
tgccagcgcc	tgaatgcaat	ctccggctac	ttgccagtac	agccggtaga	cacgctgagg	2040
atcctgagtc	tccagtcacc	ccaggaacac	caacgccgcc	agcagccgca	gcaggagcag	2100
cagcaagagg	aggaccgaga	agagaacccg	agagccggtc	tggaccctcc	ggtggcggag	2160
gaggaggagt	agctgacttg	tttcccagac	tgcgccgggt	gctgactagg	tcttccagtg	2220
gacgggagag	ggggattaag	cgggagaggc	atgaggagac	tagccacaga	actgaactga	2280
ctgtcagtct	gatgagccgc	aggcgcccag	aatcgggtgtg	gtggcatgag	gtgcagtcgc	2340
aggggataga	tgaggctctg	gtgatgcatg	agaaatattc	cctagaacaa	gtcaagactt	2400
gttggttga	gcccagggat	gattgggagg	tagccatcag	gaattatgcc	aagctggctc	2460
tgaagccaga	caagaagtac	aagattacca	aactgattaa	tatcagaaat	tcctgctaca	2520
tttcaggga	tggggccgag	gtggagatca	gtaccaggga	gaggggtggc	ttcagatggt	2580
gtatgatgaa	tatgtacccg	ggggtggtgg	gcatggaggg	agtcaccttt	atgaacacga	2640
ggttcagggg	tgatgggtat	aatgggggtg	tctttatggc	caacaccaag	ctgacagtgc	2700
acggatgctc	cttctttggc	ttcaataaca	tgtgcatcga	ggcctggggc	agtgtttcag	2760
tgaggggatg	cagcttttca	gccaactgga	tgggggtcgt	gggcagaacc	aagagcaagg	2820
tgtcagtga	gaaatgcctg	ttcgagaggt	gccacctggg	ggtgatgagc	gagggcgagg	2880
ccaaagtcaa	acactgcgcc	tctaccgaga	cgggctgctt	tgtgctgatc	aagggcaatg	2940
cccaagtcaa	gcataacatg	atctgtgggg	cctcggatga	gcgcggctac	cagatgctga	3000
cctgcgccgg	tgggaacagc	catatgctgg	ccaccgtgca	tgtggcctcg	cacccccgca	3060
agacatggcc	cgagttcgag	cacaacgtca	tgacctgctg	caatgtgcac	ctgggctccc	3120
gccgaggcat	gttcatgccc	taccagtgca	acatgcaatt	tgtgaagggtg	ctgctggagc	3180
ccgatgccat	gtccagagtg	agcctgacgg	gggtgtttga	catgaatgtg	gagctgtgga	3240
aaattctgag	atatgatgaa	tccaagacca	ggtgccgggc	ctgcgaatgc	ggaggcaagc	3300
acgccaggct	tcagcccgtg	tgtgtggagg	tgacggagga	cctgcgaccc	gatcatttgg	3360
tgttgtcctg	caacgggacg	gagttcggct	ccagcgggga	agaatctgac	tagagtgagt	3420
agtgtttggg	gctgggtgtg	agcctgcatg	aggggcagaa	tgactaaaat	ctgtggtttt	3480
ctgtgtgttg	cagcagcatg	agcgggaagc	cctcctttga	gggaggggta	ttcagccctt	3540
atctgacggg	gcgtctcccc	tcctgggcgg	gagtgcgtca	gaatgtgatg	ggatccacgg	3600
tggacggccg	gcccgtgcag	cccgcgaact	cttcaaccct	gacctacgcg	accctgagct	3660
cctcgtccgt	ggacgcagct	gccgccgcag	ctgctgcttc	cgccgccagc	gccgtgcgcg	3720
gaatggccct	gggcgcgggc	tactacagct	ctctggtggc	caactcgagt	tccaccaata	3780
atcccgccag	cctgaacgag	gagaagctgc	tgctgctgat	ggcccagctc	gaggccctga	3840

cccagcgcct	gggcgagctg	acccagcagg	tggctcagct	gcaggcggag	acgcgggccc	3900
cggttgccac	ggtgaaaacc	aaataaaaaa	tgaatcaata	aataaacgga	gacggttggt	3960
gattttaaca	cagagtcttg	aatctttatt	tgatttttcg	cgcgcggtag	gccctggacc	4020
accggtctcg	atcattgagc	acccggtgga	tcttttccag	ga'cccggtag	aggtgggctt	4080
ggatggtgag	gtacatgggc	atgagcccgt	cccgggggtg	gaggtagctc	cattgcaggg	4140
cctcgtgctc	ggggatggtg	ttgtaaatca	cccagtcata	gcagggg'cg	agggcgtggt	4200
gctgcacgat	gtccttgagg	aggagactga	tggccacggg	cagccccttg	gtgtaggtgt	4260
tgacgaacct	gttgagctgg	gagggatgca	tgcgggggga	gatgagatgc	atcctggcct	4320
ggatcttgag	attggcgatg	ttcccgccca	gatcccgccg	ggggttcatg	ttgtgcagga	4380
ccaccagcac	ggtgtatccg	gtgcacttgg	ggaatttg'tc	atgcaacttg	gaagggaagg	4440
cgtgaaagaa	tttg'gagacg	cccttg'tgac	cgcccagg'tt	ttccatgcac	tcatccatga	4500
tgatggcgat	gggcccgtgg	gcggcggcct	gggcaaagac	gtttcggggg	tcggacacat	4560
cgtagttgtg	gtcctgggtg	agctcgtcat	aggccatttt	aatgaatttg	gggcggaggg	4620
tgcccgactg	ggggacgaag	gtgccctcga	tcccgggggc	gtagttgccc	tcgcagatct	4680
gcatctccca	ggccttgagc	tcggaggggg	ggatcatgtc	cacctgcggg	gcgatgaaaa	4740
aaacgg'tttc	cggggcgggg	gagatgagct	gggccgaaag	cagg'ttccgg	agcagctggg	4800
acttgccgca	accggtgggg	ccgtagatga	ccccgatgac	cggctgcagg	tggtagt'tga	4860
gggagagaca	gctgccgtcc	tcgcggagga	ggggggccac	ctcgttcatc	atctcgcgca	4920
catgcatgtt	ctcgcgcacg	agttccgccca	ggaggcgctc	gccccccagc	gagaggagct	4980
cttg'cagcga	ggcgaagt'tt	ttcagcggct	tgagtccgtc	ggccatgggc	attttggaga	5040
gggtctgttg	caagagt'tcc	agacggtccc	agagctcgg't	gatgtgctct	agggcatctc	5100
gatccagcag	acctcctcgt	ttcgcggg'tt	ggggcgactg	cgggagtagg	gcaccaggcg	5160
atggg'cg'tcc	agcgaggcca	gggtccgg'tc	cttccagg'gc	cgcagggtcc	gcgtcagcgt	5220
gg'tctccgtc	acggtgaagg	ggtgcgcgcc	gggctggggc	cttg'cgaggg	tgcgcttcag	5280
gctcatccgg	ctggtcgaga	accgctccc'g	gtcggcgccc	tg'cgcg'tcgg	ccaggtagca	5340
attgagcatg	agttcgtagt	tgagcgcctc	ggccgcgtgg	cccttggcgc	ggagcttacc	5400
ttt'ggaagtg	tgtccgcaga	cgggacagag	gagggacttg	agggcgtaga	gcttgggggc	5460
gaggaagacg	gactcggggg	cgtaggcgtc	cgcgccgcag	ctggcgcaga	cgg'tctcgca	5520
ctccacgagc	cagg'tgaggt	cggggcggtt	gggg'tcaaaa	acgagg'tt'tc	ctccgtgctt	5580
tttgatgcgt	ttcttacctc	tgg'tctccat	gagctcgtgt	ccccgctggg	tgacaaaagag	5640
gctgtccgtg	tccccgtaga	ccgact'ttat	gggccgg'tcc	tcgagcgggg	tgccgcgg'tc	5700
ctcgtcgtag	aggaacccc'g	cccactccga	gacgaaggcc	cgggtccagg	ccagcacgaa	5760
ggaggccacg	tgggaggggt	agcgg'tcgtt	gtccaccagc	gggtccacct	tctccagggt	5820
atgcaagcac	atgtccccct	cgtccacatc	caggaagg'tg	attggctt'gt	aagtgtaggc	5880

cacgtgaccg	ggggtcccgg	ccgggggggt	ataaaagggg	gcgggcccct	gctcgtcctc	5940
actgtcttcc	ggatcgctgt	ccaggagcgc	cagctgttgg	ggtaggtatt	ccctctcgaa	6000
ggcgggcatg	acctcggcac	tcaggttgtc	agtttctaga	aacgaggagg	atttgatatt	6060
gacggtgccg	ttggagacgc	ctttcatgag	cccctcgtcc	atttggtcag	aaaagacgat	6120
ctttttgttg	tcgagcttgg	tggcgaagga	gccgtagagg	gcgttggaga	gcagcttggc	6180
gatggagcgc	atggtcttgt	tcttttcctt	gtcggcgcg	tccttggcgg	cgatgttgag	6240
ctgcacgtac	tcgcgcgcca	cgcacttcca	ttcgggggaag	acggtggtga	gctcgtcggg	6300
cacgattctg	acccgccagc	cgcggttgtg	caggggtgatg	aggccacgc	tggtggccac	6360
ctcgccgcgc	aggggctcgt	tggtccagca	gaggcgccc	cccttgcgcg	agcagaaggg	6420
gggcagcggg	tccagcatga	gctcgtcggg	ggggtcggcg	tccacggtga	agatgccggg	6480
caggagctcg	gggtcgaagt	agctgatgca	ggtgcccaga	ttgtccagcg	ccgcttgcca	6540
gtcgcgcacg	gccagcgcg	gctcgtaggg	gctgaggggc	gtgccccagg	gcatgggggtg	6600
cgtgagcgcg	gaggcgtaca	tgccgcagat	gtcgtagacg	tagaggggct	cctcgaggac	6660
gccgatgtag	gtggggtagc	agcgcccccc	gcggatgctg	gcgcgcacgt	agtcgtacag	6720
ctcgtgagag	ggcgcgagga	gccccgtgcc	gaggttggag	cgttgcggct	tttcggcgcg	6780
gtagacgatc	tggcggaaga	tggcgtggga	gttgaggag	atggtgggccc	tttggaagat	6840
gttgaagtgg	gcgtggggca	ggccgaccga	gtccctgatg	aagtgggctg	aggagtcctg	6900
cagcttggcg	acgagctcgg	cggtgacgag	gacgtccagg	gcgcagtagt	cgagggctctc	6960
ttggatgatg	tcatacttga	gctggccctt	ctgcttcac	agctcgcggt	tgagaaggaa	7020
ctcttcgcgg	tccttccagt	actcttcgag	ggggaacccg	tcctgatcgg	cacggtaaga	7080
gcccaccatg	tagaactggt	tgacggcctt	gtaggcgag	cagcccttct	ccacggggag	7140
ggcgtaagct	tgcgcgccct	tgcgcaggg	ggtgtgggtg	agggcgaagg	tgctgcgcac	7200
catgaccttg	aggaactggt	gcttgaagtc	gaggtcgtcg	cagccgccct	gctcccagag	7260
ttggaagtcc	gtgcgcttct	tgtaggcggg	gttaggcaaa	gcgaaagtaa	catcgttgaa	7320
gaggatcttg	cccgcgcggg	gcatgaagtt	gcgagtgatg	cggaaaggct	ggggcacctc	7380
ggccccggtt	ttgatgacct	gggcggcgag	gacgatctcg	tcgaagccgt	tgatgttgtg	7440
cccgacgatg	tagagttcca	cgaatcgcg	gcggcccttg	acgtggggca	gcttcttgag	7500
ctcgtcgtag	gtgagctcgg	cggggtcgct	gagcccgtgc	tgctcgaggg	cccagtcggc	7560
gacgtggggg	ttggcgctga	ggaagggaagt	ccagagatcc	acggccaggg	cggtctgcaa	7620
gcggtcccgg	tactgacgga	actgttggcc	cacggccatt	ttttcggggg	tgacgcagta	7680
gaaggtgcgg	gggtcgccgt	gccagcggtc	ccacttgagc	tggagggcga	ggtcgtgggc	7740
gagctcgacg	agcggcgggt	ccccggagag	tttcatgacc	agcatgaagg	ggacgagctg	7800
cttgccgaag	gaccccatcc	aggtgtaggt	ttccacatcg	taggtgagga	agagcctttc	7860
gggtcgagga	tgcgagccga	tggggaagaa	ctggatctcc	tgccaccagt	tggaggaatg	7920
gctgttgatg	tgatggaagt	agaaatgccg	acggcgcgcc	gagcactcgt	gcttgtgttt	7980

atacaagcgt ccgcagtgct cgcaacgctg cacgggatgc acgtgctgca cgagctgtac 8040
ctgggttcct ttggcgagga atttcagtgg gcagtggagc gctggcggct gcatctcgtg 8100
ctgtactacg tcttgccat cggcgtggcc atcgtctgcc tcgatggtgg tcatgctgac 8160
gagcccgcgc gggaggcagg tccagacctc ggctcggacg ggtcggagag cgaggacgag 8220
ggcgcgcagg ccggagctgt ccagggtcct gagacgctgc ggagtcaggt cagtgggcag 8280
cggcggcgcg cggttgactt gcaggagctt ttccagggcg cgcgggaggt ccagatggta 8340
cttgatctcc acggcgccgt tgggtggctac gtccacggct tgcagggtgc cgtgcccctg 8400
gggcgccacc accgtgcccc gtttcttctt gggcgtgct tccatgtcgg tcagaagcgg 8460
cggcgaggac gcgcgccggg cggcaggggc ggctcggggc ccggaggcag gggcggcagg 8520
ggcacgtcgg cgccgcgcgc gggcagggtt tgggtactgcg cccggagaag actggcgtga 8580
gcgacgacgc gacggttgac gtcctggatc tgacgcctct gggtgaaggc cacgggacct 8640
gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtatcgtt gacggcggcc 8700
tgccgcagga tctcttgac gtcgcccag tttgtcctgg aggcgatctc ggtcatgaac 8760
tgctcgatct cctcctcctg aaggctctcc cggccggcgc gctcgacgggt ggccgcgagg 8820
tcgttggaga tgcggcccat gagctgcgag aaggcgttca tgccggcctc gttccagacg 8880
cggctgtaga ccacggctcc gtcggggctc gcgcgcgcga tgaccacctg ggcgaggttg 8940
agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc 9000
gtggtggcga tgtgtcgggt gacgaagaag tacatgatcc agcggcggag cggcatctcg 9060
ctgacgtcgc ccagggttc caagcgttcc atggcctcgt agaagtccac ggcgaagttg 9120
aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg 9180
gcgatggtgg gcgcacctc gcgctcgaag gccccggggg gtcctcttc catctcctcc 9240
tcttctcct ccactaacat ctcttctact tctcctcag gaggcgggtg cgggggaggg 9300
gccctgcgtc gccggcggcg cacgggcaga cggtcgatga agcgtcgtat ggtctccccg 9360
cgccggcgac gcatggtctc ggtgacggcg cgcccgctct cgcggggccc cagcatgaag 9420
acgccgccgc gcatctccag gtggccgccc ggggggtctc cgttgggcag ggagagggcg 9480
ctgacgatgc atcttatcaa ttgaccgta gggactccgc gcaaggacct gagcgtctcg 9540
agatccacgg gatccgaaaa ccgctgaacg aaggcttcga gccagtcgca gtcgcaaggt 9600
aggctgagcc cggtttcttg ttcttcgggt atttggtcgg gaggcgggcg ggcgatgctg 9660
ctgggtgatga agttgaagta ggcggctctg agacggcgga tgggtggcag gagcaccagg 9720
tccttgggcc cggcttgctg gatgcgcaga cggtcggcca tgccccaggc gtggtcctga 9780
cacctggcga ggtccttgta gtagtcctgc atgagccgct ccacgggcac ctctcctcgc 9840
cccgcgcggc cgtgcatgcg cgtgagcccg aaccgcgcgt gcggctggac gagcgccagg 9900
tcggcgacga cgcgctcgggt gaggatggcc tgctggatct gggtgaggggt ggtctggaag 9960
tcgtcgaagt cgacgaagcg gtggtaggct ccgggtgttga tgggtgtagga gcagttggcc 10020

atgacggacc agttgacggt ctggtggccg ggtcgcacga gctcgtggta cttgaggcgc 10080
gagtaggcgc gcgtgtcgaa gatgtagtcg ttgcaggcgc gcacgaggta ctggtatccg 10140
acgaggaagt gcggcgccg ctggcggtag agcgccatc gctcgggtggc gggggcgccg 10200
ggcgcgaggt cctcgagcat gaggcggtgg tagccgtaga tgtacctgga catccagggtg 10260
atgccggcgg cgggtggtgga ggcgcgcggg aactcgcgga cgcggttcca gatgttgccg 10320
agcggcagga agtagttcat ggtggccgcg gtctggcccc tgaggcgcgc gcagtcgtgg 10380
atgctctaga catacgggca aaaacgaaag cggtcagcgg ctcgactccg tggcctggag 10440
gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag 10500
ccgcagctaa cgtggtactg gcactcccgt ctcgacccaa gcctgctaac gaaacctcca 10560
ggatacggag gcgggtcgtt ttttggcctt ggtcgtggt catgaaaaac tagtaagcgc 10620
ggaaagcggc cgcccgcat ggctcgtgc cgtagtctgg agaaagaatc gccagggttg 10680
cgttgcggtg tgccccggtt cgagcctcag cgtcggcgc cgccgggatt ccgcggctaa 10740
cgtgggcgtg gctgccccgt cgtttccaag accccttagc cagccgactt ctccagttac 10800
ggagcgagcc cctctttttt ttctttgtgt ttttgccaga tgcatcccgt actgcggcag 10860
atgcgcccc accctccacc acaaccgccc ctaccgcagc agcagcaaca gccggcgctt 10920
ctgccccgc cccagcagca gccagccact accgcggcgg ccgccgtgag cggagccggc 10980
gttcagtatg acctggcctt ggaagagggc gaggggctgg cgcggtggg ggcgtcgtcg 11040
ccggagcggc acccgcgct gcagatgaaa agggacgctc gcgaggccta cgtgccaag 11100
cagaacctgt tcagagacag gagcgcgag gagccgagg agatgcgcgc ctcccgcttc 11160
cacgcggggc gggagctgcg gcgcggcctg gaccgaaagc gggtgctgag ggacgaggat 11220
ttcgaggcgg acgagctgac ggggatcagc cccgcgcgc cgcacgtggc cgcgccaac 11280
ctggtcacgg cgtacgagca gaccgtgaag gaggagagca acttccaaaa atccttcaac 11340
aaccacgtgc gcacgctgat cgcgcgcgag gaggtgacct tgggcctgat gcacctgtgg 11400
gacctgctgg aggccatcgt gcagaacccc acgagcaagc cgctgacggc gcagctgttt 11460
ctggtggtgc agcacagtcg ggacaacgag acgttcaggg aggcgtgct gaatatcacc 11520
gagcccgagg gccgctggct cctggacctg gtgaacattt tgcaagcat cgtggtgcag 11580
gagcgcgggc tgccgctgtc cgagaagctg gcggccatca acttctcggg gctgagtctg 11640
ggcaagtact acgctaggaa gatctacaag acccgtacg tgcccataga caaggagggtg 11700
aagatcgacg ggttttacat gcgcatgacc ctgaaagtgc tgaccctgag cgacgatctg 11760
ggggtgtacc gcaacgacag gatgcaccgc gcggtgagcg ccagccgccg gcgcgagctg 11820
agcgaccagg agctgatgca cagcctgcag cgggccctga ccggggccgg gaccgagggg 11880
gagagctact ttgacatggg cgcggaacctg cgctggcagc ccagccgccg ggccttgga 11940
gctgccggcg gttcccccta cgtggaggag gtggacgatg aggaggagga gggcgagtac 12000
ctggaagact gatggcgcgaa ccgtatTTTT gctagatgca gcaacagcca ccgccgcgc 12060
ctcctgatcc cgcatgcgg gcggcgctgc agagccagcc gtccggcatt aactcctcgg 12120

acgattggac ccaggccatg caacgcatca tggcgctgac gacccgcaat cccgaagcct 12180
ttagacagca gcctcaggcc aaccggctct cggccatcct ggaggccgtg gtgccctcgc 12240
gctcgaaccc cagcgacgag aagggtgctg ccatcgtgaa cgcgctggtg gagaacaagg 12300
ccatccgcgg tgacgaggcc gggctggtgt acaacgcgct gctggagcgc gtggcccgcct 12360
acaacagcac caacgtgcag acgaacctgg accgcatggt gaccgacgtg cgcgaggcgg 12420
tgtcgcagcg cgagcggttc caccgcgagt cgaacctggg ctccatggtg gcgctgaacg 12480
ccttcctgag cagcgagccc gccaacgtgc cccggggcca ggaggactac accaacttca 12540
tcagcgcgct gcggctgatg gtggccgagg tgccccagag cgaggtgtac cagtcggggc 12600
cggactactt cttccagacc agtcgccagg gcttgacgac cgtgaacctg agccaggctt 12660
tcaagaactt gcagggactg tggggcgctg agggccccgt cggggaccgc gcgacggtgt 12720
cgagcctgct gacgccgaac tcgcgccctg tgctgctgct ggtggcgccc ttcacggaca 12780
gcggcagcgt gagccgcgac tcgtacctg gctacctgct taacctgtac cgcgaggcca 12840
tcggacaggc gcacgtggac gagcagacct accaggagat caccacgtg agccgcgcgc 12900
tgggcccagga ggacccgggc aacctggagg ccaccctgaa cttcctgctg accaaccggt 12960
cgcagaagat cccgccccag tacgcgctga gcaccgagga ggagcgcac ctgcgctacg 13020
tgacgacgag cgtggggctg ttcctgatgc aggagggggc cacgcccagc gcggcgctcg 13080
acatgaccgc gcgcaacatg gagcccagca tgtacgcccg caaccgcccg ttcacataa 13140
agctgatgga ctacttgcac cgggcggccg ccatgaactc ggactacttt accaacgcca 13200
tcttgaaccc gactggctc ccgccgcccg ggttctacac gggcgagtac gacatgcccg 13260
acccaacga cgggttctg tgggacgacg tggacagcag cgtgttctcg ccgcgtccag 13320
gaaccaatgc cgtgtggaag aaagagggcg gggaccggcg gccgtcctcg gcgctgtccg 13380
gtcgcgcggg tgctgccgcg gcggtgcccg agggccgacg ccccttcccg agcctgccct 13440
tttcgctgaa cagcgtgcgc agcagcgagc tgggtcggct gacgcgaccg cgcctgctgg 13500
gcgaggagga gtacctgaac gactccttgt tgaggccga gcgcgagaag aacttcccca 13560
ataacgggat agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgcacgagc 13620
acagggacga gccccgagct agcagcgagc gacccgtag acgccagcgg cacgacaggc 13680
agcggggact ggtgtgggac gatgaggatt ccgccgacga cagcagcgtg ttggacttgg 13740
gtgggagtgg tggtaacccg ttcgctcacc tgcgccccg tatcggggcg ctgatgtaag 13800
aatctgaaaa aataaaagac ggtactcacc aaggccatgg cgaccagcgt gcgttcttct 13860
ctgttggttg tagtagtatg atgaggcgcg tgtacccgga gggctcctct ccctcgtagc 13920
agagcgtgat gcagcaggcg gtggcgggcg cgatgcagcc cccgctggag gcgccttacg 13980
tgcccccgcg gtacctggcg cctacggagg ggcggaacag cattcgttac tcggagctgg 14040
cacccttgta cgataccacc cggttgtagc tggtggacaa caagtcggca gacatcgctt 14100
cgctgaacta ccagaacgac cacagcaact tcctgaccac cgtggtgcag aacaacgatt 14160

tcacccccac ggaggccagc acccagacca tcaactttga cgagcgctcg cgggtggggcg 14220
gccagctgaa aaccatcatg cacaccaaca tgcccaacgt gaacgagttc atgtacagca 14280
acaagttcaa ggcgcgggtg atggtctcgc gcaagacccc caacgggggtg gatgatgatt 14340
atgatggtag tcaggacgag ctgacctacg agtggttgga gtttgagctg cccgagggca 14400
acttctcggg gaccatgacc atcgatctga tgaacaacgc catcatcgac aactacttgg 14460
cgggtggggcg gcagaacggg gtgctggaga gcgacatcgg cgtgaagttc gacacgcgca 14520
acttccggct gggctgggac cccgtgaccg agctggtgat gccgggctg tacaccaacg 14580
aggccttcca ccccgacatc gtcctgctgc ccggtgcgg cgtggacttc accgagagcc 14640
gcctcagcaa cctgctgggc atccgcaagc ggcagccctt ccaggagggc ttccagatcc 14700
tgtacgagga cctggagggg ggcaacatcc ccgcgctctt ggatgtcgaa gcctacgaga 14760
aaagcaagga ggatagcacc gccgcggcga ccgcagccgt ggccaccgcc tctaccgagg 14820
tgcggggcca taattttgct agcgtcgcgg cagcgccga ggcggctgaa accgaaagta 14880
agatagtcac ccagccggtg gagaaggaca gcaaggacag gagctacaac gtgctcgcgg 14940
acaagaaaaa caccgcctac cgcagctggt acctggccta caactacggc gaccccgaga 15000
agggcgctgc ctcttgagc ctgctacca cctcgacgt cacctgcggc gtggagcaag 15060
tctactggtc gctgcccagc atgatgcaag acccggtcac cttccgctcc acgctcaag 15120
ttagcaacta cccggtggtg ggcgcggagc tcctgcccgt ctactccaag agcttcttca 15180
acgagcaggc cgtctactcg cagcagctgc gcgccttcac ctgctcacg cacgtcttca 15240
accgcttccc cgagaaccag atcctcgtcc gcccgcccgc gccaccatt accaccgtca 15300
gtgaaaacgt tcctgctctc acagatcacg ggaccctgcc gctgcgcagc agtatccggg 15360
gagtccagcg cgtgaccgtc actgacgcca gacgccgcac ctgcccctac gtctacaagg 15420
ccctgggctg agtcgcgccg cgcgtcctct cgagccgcac cttctaaaaa atgtccattc 15480
tcctctcgcc cagtaataac accggttggg gcctgcgcgc gccagcaag atgtacggag 15540
gcgctcgcca acgctccacg caacaccccg tgcgcgtgcg cgggcacttc cgcgtccct 15600
ggggcgccct caagggccgc gtgcgctcgc gcaccaccgt cgacgacgtg atcgaccagg 15660
tggtggccga cgcgcgcaac tacacgcccg ccgcccgcgc cgtctccacc gtggacgccg 15720
tcctcgacag cgtggtggcc gacgcgcgcc ggtacgcccg caccaagagc cggcggcggc 15780
gcatcgcccg gcggcaccgg agcacccccg ccatgcgcgc ggcgcgagcc ttgctgcgca 15840
gggcccaggc cacgggacgc agggccatgc tcagggcggc cagacgcgcg gcctccggca 15900
gcagcagcgc cggcaggacc cgcagacgcg cggccacggc ggcggcggcg gccatcgcca 15960
gcatgtcccg cccgcggcgc ggcaacgtgt actgggtgcg cgacgccgcc accggtgtgc 16020
gcgtgcccgt gcgcacccgc cccctcgcga cttgaagatg ctgacttcgc gatgttgatg 16080
tgtcccagcg gcgaggagga tgtccaagcg caaatacaag gaagagatgc tccagggtcat 16140
cgcgcctgag atctacggcc ccgcggcgcc ggtgaaggag gaaagaaagc cccgcaaact 16200
gaagcgggtc aaaaaggaca aaaaggagga ggaagatgac ggactgggtg agtttgtgcg 16260

cgagttcgcc ccccggcggc gcgtgcagtg gcgcggggcgg aaagtgaac cgggtgctgcg 16320
gccccggcacc acggtggtct tcacgcccgg cgagcggtcc ggctccgcct ccaagcgctc 16380
ctacgacgag gtgtacgggg acgaggacat cctcgagcag gcggtcgagc gtctgggcca 16440
gtttgcgtac ggcaagcgca gccgccccgc gcccttgaaa gaggaggcgg tgtccatccc 16500
gctggaccac ggcaacccca cgccgagcct gaagccggtg accctgcagc aggtgctacc 16560
gagcgcgggc cgcgccggg gcttcaagcg cgagggcggc gaggatctgt acccgaccat 16620
gcagctgatg gtgccaagc gccagaagct ggaggacgtg ctggagcaca tgaaggtgga 16680
ccccgaggtg cagcccagg tcaaggtgcg gcccatcaag caggtggccc cgggcctggg 16740
cgtgcagacc gtggacatca agatccccac ggagcccatg gaaacgcaga ccgagcccgt 16800
gaagcccagc accagcacca tggaggtgca gacggatccc tggatgccag caccagcttc 16860
caccagcact cgccgaagac gcaagtacgg cgcgcccagc ctgctgatgc ccaactacgc 16920
gctgcacct tccatcatcc ccacgccggg ctaccgcggc acgcgcttct accgcggcta 16980
caccagcagc cgccggcgca agaccaccac ccgccggcgt cgtcgagcc gccgcagcag 17040
caccgcgact tccgccttgg tgcggagagt gtatcgagc gggcgcgagc ctctgaccct 17100
gccgcgcg cgctaccacc cgagcatcgc catttaacta ccgcctccta cttgcagata 17160
tggccctcac atgccgcctc cgctcccca ttacgggcta ccgaggaaga aagccgcgcc 17220
gtagaaggct gacggggaac gggctgcgtc gccatcacca ccggcgggcg cgcgccatca 17280
gcaagcggtt ggggggaggc ttcctgcccg cgctgatccc catcatcgcc gcggcgatcg 17340
gggcgatccc cgccatagct tccgtggcgg tgcaggcctc tcagcgccac tgagacacaa 17400
aaaagcatgg atttgtaata aaaaaaaaaa tggactgacg ctcttggtcc tgtgatgtgt 17460
gttttttagat ggaagacatc aatttttcgt ccctggcacc gcgacacggc acgcggccgt 17520
ttatgggcac ctggagcgac atcggaaca gccaaactgaa cggggggcgcc ttcaattgga 17580
gcagtctctg gagcgggctt aagaatttcg ggtccacgct caaacctat ggcaacaagg 17640
cgtggaacag cagcacaggg caggcgctga gggaaaagct gaaagaacag aacttccagc 17700
agaaggtggt tgatggcctg gcctcaggca tcaacggggt ggttgacctg gccaaccagg 17760
ccgtgcagaa acagatcaac agccgcctgg acgcggtccc gcccggggg tccgtggaga 17820
tgccccaggt ggaggaggag ctgcctcccc tggacaagcg cggcgacaag cgaccgcgtc 17880
ccgacgcgga ggagacgctg ctgacgcaca cggacgagcc gccccgtac gaggaggcgg 17940
tgaaactggg cctgcccacc acgcggcccg tggcgctctt ggccaccgga gtgctgaaac 18000
ccagcagcag ccagcccgcg accctggact tgcctccgcc tcgcccctcc acagtggcta 18060
agcccctgcc gccggtggcc gtcgctcgc gcgccccccg aggccgccc caggcgaact 18120
ggcagagcac tctgaacagc atcgtgggtc tgggagtgca gagtgtgaag cgccgccgct 18180
gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt atgtccgccg 18240
accagaagga ggagtgtgaa gaggcgcgtc gccgagttgc aagatggcca ccccatcgat 18300

gctgccccag tggggtgtaca tgcacatcgc cggacaggac gcttcggagt acctgagtc 18360
gggtctggtg cagttcgcgc gcgccacaga cacctacttc agtctgggga acaagttag 18420
gaacccacg gtggcgccca cgcacgatgt gaccaccgac cgcagccagc ggctgacgct 18480
gcgcttcgtg cccgtggacc gcgaggacaa cacctactcg taaaaagtgc gctacacgct 18540
ggcgtgggc gacaaccgcg tgctggacat ggccagcacc tactttgaca tccgcggcgt 18600
gctggaccgg ggccctagct tcaaacccta ctctggcacc gcctacaaca gcctagctcc 18660
caagggagct cccaattcca gccagtggga gcaagcaaaa acaggcaatg ggggaactat 18720
ggaaacacac acatatggtg tggcccaat gggcggagag aatattaca aagatggtct 18780
tcaaattgga actgacgtta cagcgaatca gaataaacca atttatgccg aaaaaacatt 18840
tcaaccagaa ccgcaagtag gagaagaaaa ttggcaagaa actgaaaact tttatggcgg 18900
tagagctctt aaaaaagaca caaacatgaa accttgctat ggctcctatg ctagaccac 18960
caatgaaaaa ggaggtcaag ctaaacctaa agttggagat gatggagttc caaccaaga 19020
attcgacata gacctggctt tctttgatac tcccgggtggc accgtgaacg gtcaagacga 19080
gtataaagca gacattgtca tgtataccga aaacacgtat ttgaaaactc cagacacgca 19140
tgtggtatac aaaccaggca aggatgatgc aagttctgaa attaacctgg ttcagcagtc 19200
tatgccaac agaccaact acattgggtt cagggacaac tttatcggtc ttatgtacta 19260
caacagcact ggcaatatgg gtgtgcttgc tggtcaggcc tcccagctga atgctgtggt 19320
tgatttgcaa gacagaaaca ccgagctgtc ctaccagctc ttgcttgact ctttgggtga 19380
cagaacccgg tatttcagta tgtggaacca ggcggtggac agttatgacc ccgatgtgcg 19440
catcatcgaa aaccatggtg tggaggatga attgccaaac tattgcttcc ccttggacgg 19500
ctctggcact aacgccgcat accaagggtg gaaagtaaaa gatggtcaag atggtgatgt 19560
tgagagtga tgggaaaatg acgatactgt tgcagctcga aatcaattat gtaaaggtaa 19620
cattttcgcc atggagatta atctccaggc taacctgtgg agaagtttcc tctactcgaa 19680
cgtggccctg tacctgcccc actcctacaa gtacacgccg accaacgtca cgctgccgac 19740
caacaccaac acctacgatt acatgaatgg cagagtgaac cctccctcgc tggtagacgc 19800
ctacctcaac atcggggcgc gctggctcgt ggaccccatg gacaacgtca accccttcaa 19860
ccaccaccgc aacgcgggccc tgcgctaccg ctccatgctc ctgggcaacg ggcgctacgt 19920
gcccttccac atccagggtgc cccaaaagtt tttcgccatc aagagcctcc tgctcctgcc 19980
cgggtcctac acctacgagt ggaacttccg caaggacgtc aacatgatcc tgcagagctc 20040
cctaggcaac gacctgcgca cggacggggc ctccatcgcc ttcaccagca tcaacctcta 20100
cgccaccttc ttccccatgg cgcacaacac cgcctccacg ctcgaggcca tgctgcgcaa 20160
cgacaccaac gaccagtcct tcaacgacta cctctcggcg gccaacatgc tctaccccat 20220
cccggccaac gccaccaacg tgcccatctc catcccctcg cgcaactggg ccgccttccg 20280
cggatggtcc ttcacgcgcc tgaagacccg cgagacgccc tcgctcggct ccgggttcga 20340
cccctacttc gtctactcgg gctccatccc ctacctagac ggcaccttct acctcaacca 20400

caccttcaag aaggtctcca tcaccttcca ctctccgctc agctggcccg gcaacgaccg 20460
cctcctgacg cccaacgagt tcgaaatcaa gcgcaccgctc gacggagagg gatacaacgt 20520
ggcccagtg c aacatgacca aggactgggtt cctgggtccag atgctggccc actacaacat 20580
cggctaccag ggcttctacg tgcccagagg ctacaaggac cgcatgtact ccttcttccg 20640
caacttccag cccatgagcc gccaggctcgt ggacgaggctc aactacaagg actaccaggc 20700
cgtcaccctg gcctaccagc acaacaactc gggcttccgctc ggctacctcg cgcccaccat 20760
gcgccagggc cagccctacc ccgccaacta cccctacccg ctcatcggca agagcgccgt 20820
cgccagcgtc acccagaaaa agttcctctg cgaccgggtc atgtggcgca tcccccttctc 20880
cagcaacttc atgtccatgg gcgcgctcac cgacctcggc cagaacatgc tctacgccaa 20940
ctccgcccac gcgctagaca tgaatttcca agtcgacccc atggatgagt ccacccttct 21000
ctatgttgctc ttcgaagtct tcgacgtcgt ccgagtgcac cagccccacc gcggcgtcat 21060
cgaagccgtc tacctgcgca cgcccttctc ggccggcaac gccaccacct aagccgctct 21120
tgcttcttgc aagatgacgg cggggtccgg cgagcaggag ctcaggggca tcctccgcga 21180
cctgggctgc gggccctgct tcctgggcac cttcgacaag cgcttccctg gattcatggc 21240
cccgacacaag ctggcctgcg ccacgtgaa cagggccggc cgcgagaccg ggggcgagca 21300
ctggctggcc ttcgcctgga acccgcgctc ccacacatgc tacctcttcg accccttcgg 21360
gttctcggac gagcgctca agcagatcta ccagttcgag tacgagggcc tgctgcgtcg 21420
cagcgccctg gccaccgagg accgctgcgt caccctggaa aagtccaccc agaccgtgca 21480
gggtccgcgc tcggccgcct gcgggtctt ctgctgcatg ttcctgcacg ccttcgtgca 21540
ctggccccgac cgccccatgg acaagaacct caccatgaac ttactgacgg gggtgcccaa 21600
cggcagtctc cagtcgcccc aggtggaacc caccctgcgc cgcaaccagg aagcgctcta 21660
ccgcttctc aatgccact ccgcctactt tcgctccac cgcgcgcgca tcgagaaggc 21720
caccgccttc gaccgcatga atcaagacat gtaaaaaacc ggtgtgtgta tgtgaatgct 21780
ttattcataa taaacagcac atgtttatgc caccttctct gaggtcttga ctttatttag 21840
aaatcgaagg ggttctgccg gctctcggca tggcccgcgg gcagggatac gttgcggaac 21900
tggtacttgg gcagccactt gaactcgggg atcagcagct tgggcacggg gaggtcgggg 21960
aacgagtcgc tccacagctt gcgcgtgagt tgcagggcgc ccagcaggctc gggcgcgagg 22020
atcttgaaaat cgcagttggg acccgcgctt tgcgcgcgag agttgcggta cacgggggtt 22080
cagcactgga acaccatcag ggccgggtgc ttcacgcttg ccagcaccgt cgcgctcggtg 22140
atgccctcca cgtccagatc ctgggcgttg gccatcccga aggggggtcat cttgcaggctc 22200
tgccgccccca tgctgggcac gcagccgggc ttgtgggttg aatcgcagt cagggggatc 22260
agcatcatct gggcctgctc ggagctcatg cccgggtaca tggccttcat gaaagcctcc 22320
agctggcgga aggcctgctg cgcttgccg ccctcggtga agaagacccc gcaggacttg 22380
ctagagaact ggttggtggc gcagccggcg tcgtgcacgc agcagcgcg gtcgttggtg 22440

gccagctgca ccacgctgcg cccccagcgg ttctgggtga tcttggcccg gttgggggttc 22500
 tccttcagcg cgcgctgccc gttctcgctc gccacatcca tctcgatagt gtgctccttc 22560
 tggatcatca cgggtcccgtg caggcaccgc agcttgccct cggcttcggt gcagccgtgc 22620
 agccacagcg cgcagccggt gcactcccag ttcttgtggg cgatctggga gtgcgagtgc 22680
 acgaagccct gcaggaagcg gcccattcatc gcggtcaggg tcttgttgct ggtgaaggctc 22740
 agcgggatgc cgcggtgctc ctcgttcaca tacagggtggc agatgcggcg gtacacctcg 22800
 ccctgctcgg gcatcagctg gaaggcggac ttcaggctgc tctccacgcg gtaccgggtcc 22860
 atcagcagcg tcatcacttc catgcccttc tcccaggccg aaacgatcgg caggctcagg 22920
 gggttcttca ccgccattgt catcttagtc gccgccgccg aggtcagggg gtcgttctcg 22980
 tccagggtct caaacactcg cttgccgtcc ttctcgatga tgcgcacggg gggaaagctg 23040
 aagcccacgg ccgccagctc ctctcggcc tgcctttcgt cctcgctgtc ctggctgatg 23100
 tcttgcaaag gcacatgctt ggtcttgcg ggtttctttt tgggcggcag aggcggcggc 23160
 gatgtgctgg gagagcgcga gttctcgttc accacgacta tttcttcttc ttggccgtcg 23220
 tccgagacca cgcggcggta ggcatgcctc ttctggggca gaggcggagg cgacgggctc 23280
 tcgcggttcg gcgggcggct ggcagagccc cttccgcgtt cgggggtgcg ctctggcgg 23340
 cgctgctctg actgacttcc tccgcggccg gccattgtgt tctcctaggg agcaacaaca 23400
 agcatggaga ctacgccatc gtcgccaaca tcgccatctg cccccgccgc caccgccgac 23460
 gagaaccagc agcagaatga aagcttaacc gccccgccgc ccagccccac ctccgacgcc 23520
 gcggcccccag acatgcaaga gatggaggaa tccatcgaga ttgacctggg ctacgtgacg 23580
 cccgcggagc acgaggagga gctggcagcg cgcttttcag ccccggaaga gaaccaccaa 23640
 gagcagccag agcaggaagc agagaacgag cagaaccagg ctgggcacga gcatggcgac 23700
 tacttgagcg gggcagagga cgtgctcatc aagcatctgg cccgccaatg catcatcgtc 23760
 aaggacgcgc tgctcgaccg cgccgaggtg cccctcagcg tggcggagct cagccgcgcc 23820
 tacgagcgca acctcttctc gccgcgcgtg cccccaaagc gccagcccaa cggcacctgt 23880
 gagcccaacc cgcgcctcaa cttctaccg gtcttcgcgg tgcccaggc cctggccacc 23940
 taccacctct ttttcaagaa ccaaaggatc cccgtctcct gccgcgcaa ccgcacccgc 24000
 gccgacgcc tgctcaacct gggccccggc gccgcctac ctgatatcac ctcttgga 24060
 gaggttccca agatcttcga gggctctgggc agcgacgaga ctcgggccgc gaacgctctg 24120
 caaggaagcg gagaggagca tgagcaccac agcgccctgg tggagttgga aggcgacaac 24180
 gcgcgcctgg cggtcctcaa gcgcacggctc gagctgacct acttcgccta cccggcgctc 24240
 aacctgcccc ccaaggctcat gagcgccgtc atggaccagg tgctcatcaa gcgcgcctcg 24300
 cccctctcgg aggaggagat gcaggacccc gagagtctcg acgagggcaa gcccgtggctc 24360
 agcgacgagc agctggcgcg ctggctggga gcgagtagca cccccagag cctggaagag 24420
 cggcgcaagc tcatgatggc cgtggctcctg gtgaccgtgg agctggagtg tctgcgccgc 24480
 ttctttgccc acgcggagac cctgcgcaag gtcgaggaga acctgcacta cctcttcagg 24540

cacgggttcg tgcgccaggc ctgcaagatc tccaacgtgg agctgaccaa cctgggtctcc 24600
tacatgggca tcctgcacga gaaccgcctg gggcaaaacg tgctgcacac caccctgcgc 24660
ggggaggccc gccgcgacta catccgcgac tgcgtctacc tgtacctctg ccacacctgg 24720
cagacgggca tgggcgtgtg gcagcagtg cctggaggagc agaacctgaa agagctctgc 24780
aagctcctgc agaagaacct caaggccctg tggaccgggt tcgacgagcg taccaccgcc 24840
tcggacctgg ccgacctcat cttccccgag cgcttcgagg tgacgctgcg caacgggctg 24900
cccgacttta tgagccaaag catgttgcaa aactttcgct ctttcatcct cgaacgctcc 24960
gggatcctgc ccgccacctg ctccgcgctg ccctcggact tcgtgccgct gaccttccgc 25020
gagtgcctcc cgccgctctg gagccactgc tacttgctgc gcctggccaa ctacctggcc 25080
taccactcgg acgtgatcga ggacgtcagc ggcgagggtc tgctggagtg cactgcccgc 25140
tgcaacctct gcacgccgca ccgctccctg gcctgcaacc ccagctgct gagcgagacc 25200
cagatcatcg gcaccttcga gttgcaaggc cccggcgacg gcgagggcaa ggggggtctg 25260
aaactacccc cggggctgtg gacctcggcc tacttgcgca agttcgtgcc cgaggactac 25320
catcccttcg agatcagggt ctacgaggac caatcccagc cgccaaggc cgagctgtcg 25380
gcctgcgtca taccacaggg ggccatcctg gcccaattgc aagccatcca gaaatcccgc 25440
caagaatttc tgctgaaaaa gggccacggg gtctacttgg acccccagac cggagaggag 25500
ctcaaccca gcttcccca ggatgccccg aggaagcagc aagaagctga aagtggagct 25560
gccgccgccc gaggatttgg aggaagactg ggagagcagt caggcagagg aggaggagat 25620
ggaagactgg gacagcactc aggcagagga ggacagcctg caagacagtc tggaggagga 25680
agacgaggtg gaggaggcag aggaagaagc agccgccgccc agaccgtcgt cctcggcgga 25740
gaaagcaagc agcacggata ccatctccgc tccgggtcgg ggtcgcggcg gccgggcca 25800
cagtaggttg gacgagaccg ggcgcttccc gaacccacc acccagaccg gtaagaagga 25860
gcggcagggg tacaagtctt ggcgggggca caaaaacgcc atcgtctcct gcttgcaagc 25920
ctgcgggggc aacatctcct taccggcgct ctacctgctc ttccaccgcg ggggtgaactt 25980
cccccgcaac atcttgcatc actaccgtca cctccacagc ccctactact gtttccaaga 26040
agaggcagaa acccagcagc agcagaaaac cagcggcagc agcagctaga aaatccacag 26100
cggcggcagg tggactgagg atcgcggcga acgagccggc gcagaccggg gagctgagga 26160
accggatctt tcccaccctc tatgccatct tccagcagag tcgggggcag gagcaggaa 26220
tgaaagtcaa gaaccgttct ctgcgctcgc tcaccgcag ttgtctgtat cacaagagcg 26280
aagaccaact tcagcgcact ctcgaggacg ccgaggctct cttcaacaag tactgcgcgc 26340
tactcttaa agagtagccc gcgcccggc acacacggaa aaaggcggga attacgtcac 26400
cacctgcgcc cttcgcccga ccatcatgag caaagagatt cccacgcctt acatgtggag 26460
ctaccagccc cagatgggccc tggccgcccg cgccgcccag gactactcca cccgcatgaa 26520
ctggctcagt gccgggcccg cgatgatctc acgggtgaat gacatccgcg cccaccgaaa 26580

ccagatactc ctagaacagt cagcgatcac cgccacgccc cgccatcacc ttaatccgcg 26640
 taattggccc gccgccttg tgtaccagga aattccccag cccacgaccg tactacttcc 26700
 gcgagacgcc caggccgaag tccagctgac taactcaggt gtccagctgg ccggcggcgc 26760
 cgccctgtgt cgtcaccgcc ccgctcaggg tataaagcgg ctggtgatcc gaggcagagg 26820
 cacacagctc aacgacgagg tggtagactc ttcgctgggt ctgcgacctg acggagtctt 26880
 ccaactcgcc ggatcgggga gatcttcctt cacgcctcgt caggccgtcc tgactttgga 26940
 gagttcgtcc tcgcagcccc gctcgggcgg catcggcact ctccagttcg tggaggagtt 27000
 cactccctcg gtctacttca accccttctc cggctcccc ggccactacc cggacgagtt 27060
 catcccgaac ttcgacgcca tcagcgagtc ggtggacggc tacgattgaa tgtcccatgg 27120
 tggcgcagct gacctagctc ggcttcgaca cctggaccac tgccgccgct tccgctgctt 27180
 cgctcgggat ctgcgcgagt ttgcctactt tgagctgccc gaggagcacc ctcaggggcc 27240
 agccacgga gtgcggatca tcgtcgaagg gggcctcgac tccacctgc ttcggatctt 27300
 cagccagcga ccgatcctgg tcgagcgcg acaaggacag acccttctta ctttgtactg 27360
 catctgcaac caccgcggcc tgcatgaaag tctttgttgt ctgctgtgta ctgagtataa 27420
 taaaagctga gatcagcgac tactccggac tcgattgtgg tgttcctgct atcaaccggt 27480
 ccctgttctt caccgggaac gagaccgagc tccagctcca gtgtaagccc cacaagaagt 27540
 acctcacctg gctgttccag ggctccccga tcgccgttgt caaccactgc gacaacgacg 27600
 gagtcctgct gagcggccct gccaacctta cttttccac ccgcagaagc aagctccagc 27660
 tcttccaacc ctctctcccc gggacctatc agtgcgcttc aggacctgc catcacacct 27720
 tccacctgat cccgaatacc acagcgccgc tccccgctac taacaaccaa actaccacc 27780
 aacgccaccg tcgcgacctt tcctctgaat ctaataccac taccggaggt gagctccgag 27840
 gtcgaccaac ctctgggatt tactacggcc cctgggaggt ggtgggggta atagcgctag 27900
 gcctagttgc ggggtgggctt ttgggttctt gctacctata cctcccttgc tgttcgtact 27960
 tagtggtgct gtgttgctgg ttttaagaaat ggggaagatc accctagtga gctgcggtgc 28020
 gctgggtggc gtgttgcttt cgattgtggg actgggcggc gcggctgtag tgaaggagaa 28080
 ggccgatccc tgcttgcat tcaatcccaa caaatgccag ctgagttttc agcccgatgg 28140
 caatcgggtgc gcggtactga tcaagtgcgg atgggaatgc gagaacgtga gaatcgagta 28200
 caataacaag actcgggaaca atactctcgc gtccgtgtgg cagcccgggg accccgagtg 28260
 gtacaccgtc tctgtccccg gtgctgacgg ctccccgcgc accgtgaata atactttcat 28320
 ttttgccac atgtgcaaca cggctcatgt gatgagcaag cagtacgata tgtggcccc 28380
 cacgaaggag aacatcgtgg tcttctccat cgcttacagc ctgtgcacgg cgctaatac 28440
 cgctatcgtg tgcttgagca ttcacatgct catcgctatt cggcccagaa ataatgccga 28500
 gaaagagaaa cagccataac acgtttttt acacaccttg tttttacaga caatgcgtct 28560
 gttaaatttt ttaaacattg tgctcagtat tgcttatgcc tctgggttatg caaacataca 28620
 gaaaaccctt tatgtaggat ctgatggtac actagagggt acccaatcac aagccaaggt 28680

tgcattggtat ttttatagaa ccaacactga tccagttaaa ctttgtaagg gtgaattgcc 28740
 gcgtacacat aaaactccac ttacatttag ttgcagcaat aataatctta cacttttttc 28800
 aattacaaaa caatatactg gtacttatta cagtacaaac tttcatacag gacaagataa 28860
 atattatact gttaaggtag aaaatcctac cactcctaga actaccacca ccaccactac 28920
 tgcaaagccc actgtgaaaa ctacaactag gaccaccaca actacagaaa ccaccaccag 28980
 cacaacactt gctgcaacta cacacacaca cactaagcta accttacaga ccactaatga 29040
 tttgatcgcc ctgctgcaaa aggggggataa cagcaccact tccaatgagg agatacccaa 29100
 atccatgatt ggcattattg ttgctgtagt ggtgtgcatg ttgatcatcg ccttgtgcat 29160
 ggtgtactat gccttctgct acagaaagca cagactgaac gacaagctgg aacacttact 29220
 aagtgttgaa ttttaatttt ttagaaccat gaagatccta ggccttttta gtttttctat 29280
 cattacctct gctctttgtg aatcagtga tagagatgtt actattacca ctggttctaa 29340
 ttatacactg aaagggccac cctcaggtat gctttcgtgg tattgctatt ttggaactga 29400
 cactgatcaa actgaattat gcaattttca aaaaggcaaa acctcaaact ctaaaatctc 29460
 taattatcaa tgcaatggca ctgatctgat actactcaat gtcacgaaag catatggtgg 29520
 cagttattat tgccctggac aaaacactga agaaatgatt ttttataaag tggaagtggg 29580
 tgatcccact acaccacca ccaccacaac tattcatacc acacacacag aacaaacacc 29640
 agaggcaaca gaagcagagt tggccttcca gggtcacgga gattcctttg ctgtcaatac 29700
 ccctacacc gatcagcggg gtccggggcc gctagtcagc ggcattgtcg gtgtgctttc 29760
 gggattagca gtcataatca tctgcatgtt ctttttgct tgctgctata gaaggcttta 29820
 ccgacaaaaa tcagaccac tgctgaacct ctatgtttaa ttttttccag agccatgaag 29880
 gcagttagcg ctctagtttt ttgttctttg attggcattg tttttaatag taaaattacc 29940
 agagttagct ttattaaaca tgttaatgta actgaaggag ataacatcac actagcaggt 30000
 gtagaagggt ctcaaaacac cacctggaca aaataccatc taggatggag agatatttgc 30060
 acctggaatg taacttatta ttgcatagga gttaatctta ccattgttaa cgctaaccaa 30120
 tctcagaatg ggttaattaa aggacagagt gttagtgtga ccagtgatgg gtactatacc 30180
 cagcatagtt ttaactacaa cattactgtc ataccactgc ctacgcctag ccacactagc 30240
 actaccacac agacaaccac atacagtaca tcaaatcagc ctaccaccac tacagcagca 30300
 gaggttgcca gctcgtctgg ggtccgagtg gcatttttga tgttggtccc atctagcagt 30360
 cccactgcta gtaccaatga gcagactact gaatttttgt ccactgtcga gagccacacc 30420
 acagctacct ccagtgctt ctctagcacc gccaatctct cctcgttttc ctctacacca 30480
 atcagccccg ctactactcc tagccccgt cctcttccca ctcccctgaa gcaaacagac 30540
 ggcggcatgc aatggcagat caccctgctc attgtgatcg gggttggtcat cctggccgtg 30600
 ttgctctact acatcttctg ccgccgcatt cccaacgcgc accgcaagcc ggcctacaag 30660
 cccatcggtta tcgggcagcc ggagccgctt caggtggaag ggggtctaag gaatcttctc 30720

ttctcttttta cagtatggtg attgaactat gattcctaga caattcttga tcactattct 30780
tatctgcctc ctccaagtct gtgccaccct cgctctggtg gccaacgcca gtccagactg 30840
tattgggccc ttcgcctcct acgtgctctt tgccttcgtc acctgcatct gctgctgtag 30900
catagtctgc ctgcttatca ccttcttcca gttcattgac tggatctttg tgcgcatcgc 30960
ctacctgcgc caccaccccc agtaccgcga ccagcgagtg gcgcagctgc tcaggctcct 31020
ctgataagca tgcgggctct gctacttctc gcgcttctgc tgtagtgct cccccgtccc 31080
gtcgaccccc ggtccccac tcagtcccc gaggaggttc gcaaatagcaa attccaagaa 31140
ccctggaaat tcctcaaag ctaccgcca aaatcagaca tgcatcccag ctggatcatg 31200
atcattggga tcgtgaacat tctggcctgc accctcatct cttttgtgat ttaccctgc 31260
tttgactttg gttggaactc gccagaggcg ctctatctcc cgcctgaacc tgacacacca 31320
ccacagcagc aacctcaggc acacgcacta ccaccaccac agcctaggcc acaatacatg 31380
cccatattag actatgaggc cgagccacag cgacctatgc tccccgctat tagttacttc 31440
aatctaaccg gcggagatga ctgaccact ggccaataac aacgtcaacg accttctcct 31500
ggacatggac ggccgcgcct cggagcagcg actcgcccaa cttcgcatc gtcagcagca 31560
ggagagagcc gtcaaggagc tgcaggacgg catagccatc caccagtga agagaggcat 31620
cttctgcctg gtgaaacagg ccaagatctc ctacgaggtc acccagaccg accatcgcct 31680
ctcctacgag ctctgcagc agcgccagaa gttcacctgc ctggtcggag tcaaccccat 31740
cgtcatcacc cagcagtcgg gcgataccaa ggggtgcac cactgctcct gcgactcccc 31800
cgactgcgtc cacactctga tcaagaccct ctgcggcctc cgcgacctcc tccccatgaa 31860
ctaatacccc ctttatccag tgaataaag atcatattga tgatgattta aataaaaaaa 31920
ataatcattt gatttgaaat aaagatacaa tcatattgat gatttgagtt taacaaaaat 31980
aaagaatcac ttacttgaaa tctgatacca ggtctctgtc catgttttct gccaacacca 32040
cctcactccc ctcttccag ctctggtact gcaggccccg gcgggctgca aacttctctc 32100
acacgctgaa ggggatgtca aattcctcct gtccctcaat cttcatttta tcttctatca 32160
gatgtccaaa aagcgcgtcc ggggtgatga tgacttcgac cccgtctacc cctacgatgc 32220
agacaacgca ccgaccgtgc cttcatcaa ccccccttc gtctcttcag atggattcca 32280
agagaagccc ctgggggtgt tgtccctgcg actggctgac cccgtcacca ccaagaacgg 32340
ggaaatcacc ctcaagctgg gagagggggt ggacctcgac tcgtcgggaa aactcatctc 32400
caacacggcc accaaggccg ccgccccctc cagtatttca aacaacacca tttcccttaa 32460
aactgctgcc ctttctaca acaacaatgg aactttaagc ctcaatgtct ccacaccatt 32520
agcagtattt cccacattta acactttagg cataagtctt ggaaacggtc ttcagacttc 32580
aaataagttg ttgactgtac aactaactca tcctcttaca ttcagctcaa atagcatcac 32640
agtaaaaaca gacaaagggc tatatattaa ctccagtggg aacagaggac ttgaggctaa 32700
tataagccta aaaagaggac tagtttttga cggtaatgct attgcaacat atattggaaa 32760
tggcttagac tatggatctt atgatagtga tggaaaaaca agaccgtaa ttaccaaata 32820

tggagcagga ttaaattttg atgctaacaa agcaatagct gtcaaactag gcacaggttt 32880
aagttttgac tccgctggtg ccttgacagc tggaaacaaa caggatgaca agctaacact 32940
ttggactacc cctgacccaa gccctaattg tcaattactt tcagacagag atgccaaatt 33000
tactctctgt cttacaaaat gcggtagtca aatactaggc actgtggcag tggcggctgt 33060
tactgtagga tcagcactaa atccaattaa tgacacagtc aaaagcgcca tagttttcct 33120
tagatttgat tccgatggtg tactcatgtc aaactcatca atggtaggtg attactggaa 33180
ctttagggag ggacagacca ctcaaagtgt agcctataca aatgctgtgg gattcatgcc 33240
aaatataggt gcatatccaa aaacccaaag taaaacacct aaaaatagca tagtcagtca 33300
ggtatattta actggagaaa ctactatgcc aatgacacta accataactt tcaatggcac 33360
tgatgaaaaa gacacaaccc cagttagcac ctactctatg acttttacat ggcatggac 33420
tggagactat aaggacaaaa atattacctt tgctaccaac tcattctctt tttcctacat 33480
cgcccaggaa taatcccacc cagcaagcca acccctttt cccaccctt tgtctatatg 33540
gaaactctga aacagaaaaa taaagttcaa gtgtttttatt gaatcaacag ttttacagga 33600
ctcgagcagt tatttttcct ccaccctccc aggacatgga atacaccacc ctctcccccc 33660
gcacagcctt gaacatctga atgccattgg tgatggacat gcttttggtc tccacgttcc 33720
acacagtttc agagcgagcc agtctcggat cggtcaggga gatgaaaccc tccgggact 33780
cccgcactctg cacctcacag ctcaacagct gaggattgtc ctcggtggtc gggatcacgg 33840
ttatctggaa gaagcagaag agcggcggtg ggaatcatag tccgcgaacg ggatcggccg 33900
gtggtgtcgc atcaggcccc gcagcagtcg ctgccgccgc cgctccgtca agctgctgct 33960
caggggggtc ggggtccaggg actccctcag catgatgcc acggccctca gcatcagtcg 34020
tctggtgctg cgggctcagc agcgcagtcg aatctcgctc aggtcactgc agtacgtgca 34080
acacaggacc accaggttgt tcaacagtcc atagttcaac acgctccagc cgaaactcat 34140
cgcggaagg atgctacca cgtggccgtc gtaccagatc ctcaggtaaa tcaagtggcg 34200
ctccctccag aagacgtgc ccatgtacat gatctccttg ggcatgtggc ggttcaccac 34260
ctcccgttac cacatcacc tctggttgaa catgcagccc cggatgatcc tgcggaacca 34320
cagggccagc accgccccgc ccgcatgca gcgaagagac cccggatccc ggcaatgaca 34380
atggaggacc caccgctcgt acccgtggat catctgggag ctgaacaagt ctatgttggc 34440
acagcacagg catatgctca tgcactctt cagcactctc agctcctcgg gggcctaac 34500
catatcccag ggcacgggga actcttgagc gacagcgaac cccgcagaac agggcaatcc 34560
tcgcacataa cttacattgt gcatggacag ggtatcgcaa tcaggcagca ccgggtgatc 34620
ctccaccaga gaagcgcggg tctcggctc ctcacagcgt ggtaaggggg ccggccgata 34680
cgggtgatgg cgggacgcgg ctgacgtgt tctcgaccgt gtcacgatgc agttgctttc 34740
ggacattttc gtacttgctg tagcagaacc tgggtccggc gctgcacacc gatcgccggc 34800
ggcggctctg gcgcttgga cgctcgggtg taaagttgta aaacagccac tctctcagac 34860

cgtgcagcag atctagggcc tcaggagtga tgaagatccc atcatgcctg atagctctga 34920
 tcacatcgac caccgtggaa tgggccaggc ccagccagat gatgcaattt tgttgggttt 34980
 cggtgacggc gggggaggga agaacaggaa gaaccatgat taacttttaa tccaaacggt 35040
 ctcggagcac ttcaaaatga aggtcacgga gatggcacct ctgcccccg ctgtgttgg 35100
 ggaaaataac agccagggtca aagggtgatac ggttctcgag atgttccacg gtggcttcca 35160
 gcaaagcctc cacgcgacac tccagaaaca agacaatagc gaaagcggga gggttctcta 35220
 attcctcaac catcatgtta cactcctgca ccatccccag ataattttca tttttccagc 35280
 cttgaatgat tcgaactagt tcctgaggta aatccaagcc agccatgata aaaagctcgc 35340
 gcagagcacc ctccaccggc attcttaagc acaccctcat aattccaaga tattctgctc 35400
 ctgggttcacc tgcagcagat tgacaagcgg aatatcaaaa tctctgccgc gatccctgag 35460
 ctctccctc agcaataact gtaagtactc tttcatatcg tctccgaaat ttttagccat 35520
 aggaccccca ggaataagag aagggcaagc cacattacag ataaaccgaa gtccccccca 35580
 gtgagcattg ccaaagttaa gattgaaata agcatgctgg ctagaccggg tgatatcttc 35640
 cagataactg gacagaaaat cgggtaagca atttttaaga aaatcaacaa aagaaaaatc 35700
 ttccagggtc acgttttaggg cctcgggaac aacgatggag taagtgaag ggggtgcgttc 35760
 cagcatgggt agtttagctga tctgtaaaaa aacaaaaaat aaaacattaa accatgctag 35820
 cctggcgaac aggtgggtaa atcggttctc ccagcaccag gcaggccacg ggggtctccg 35880
 cgcgaccctc gtaaaaattg tcgctatgat tgaaaaccat cacagagaga cgttcccgg 35940
 ggccggcggtg aatgattcga gaagaagcat acaccccggt aacattggag tccgtgagtg 36000
 aaaaaaagcg gccgaggaag caatgaggca ctacaacgct cactctcaag tccagcaaag 36060
 cgatgccatg cggatgaagc acaaaatttt cagggtgcgtg aaaaatgtaa ttactcccct 36120
 cctgcacagg cagcgaagct cccgatccct ccagatacac atacaaagcc tcagcgtcca 36180
 tagcttaccg agcggcagca gcagcggcac acaacaggcg caagagtcag agaaaagact 36240
 gagctctaac ctgtccgccc gctctctgct caatatatag cccagatct aactgacgt 36300
 aaaggccaaa gtctaaaaat acccgccaaa taatcacaca cgcccagcac acgcccagaa 36360
 accggtgaca cactcagaaa aatacgcgca cttcctcaaa cggccaaact gccgtcattt 36420
 ccgggttccc acgctacgtc atcaaaacac gactttcaaa ttccgtcgac cgttaaaaac 36480
 atcacccgcc ccgcccctaa cggtcgccc tcccgcagcc aatcaccttc ctccctcccc 36540
 aaattcaaac agctcatttg catattaacg cgcacaaaaa gtttgaggta tattattgat 36600
 gatg 36604

<210> 3
 <211> 36535
 <212> DNA
 <213> chimpanzee adenovirus serotype Pan7

<400> 3
 catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaag agctgtttga 60

atttggggag ggaggaaggt gattggccga gagacgggag accgttaggg gcggggcggg 120
 tgacgttttt aatacgtggc cgtgaggcgg agccgggtttg caagttctcg tgggaaaagt 180
 gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
 ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
 aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
 ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcacctaaat 420
 ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cgccagggtgta 480
 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagttttct 540
 cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 600
 gtaatgtttt cctggctact gggaacgaga ttctggaatt ggtggtggac gccatgatgg 660
 gtggcgaccc tcctgagccc cctaccccat ttgaggcgcc ttcgctgtac gatttgatg 720
 atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta 780
 gcgatgccgc gctgctggct gccgagcagg ctaatacggg ctctggctca gacagcgatt 840
 cctctctcca taccgagaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
 aagagctcga cctgcgctgc tatgaggaat gcttgccctc gagcgatgat gaggaggacg 960
 aggaggcgat tcgagctgca tcgaaccagg gagtgaagc tgcgggcgaa agcttttagcc 1020
 tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcac cgcatgaata 1080
 ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac cattgtgttt 1140
 acagtaagtg tgattaactt tagttgggaa ggcagagggg gactgggtgc tgactgggtt 1200
 atttatgtat atgttttttt atgtgtaggt cccgtctctg acgtagatga gacccccact 1260
 tcagagtgca tttcatcacc cccagaaatt ggcgaggaac cgcccgaaga tattattcat 1320
 agaccagtgt cagtgagagt caccgggagg agagcagctg tggagagttt ggatgacttg 1380
 ctacaggggtg gggatgaacc tttggacttg tgtaccgga aacgccccag gactaagtg 1440
 ccacacatgt gtgtttactt aaggatgatg cagtatttat aggggtgtga gtgcaataaa 1500
 atccgtgttg actttaagtg cgtgggtttat gactcagggg tggggactgt gggtatataa 1560
 gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gacggtcttg 1620
 gaagactttc accagactag acagctgcta gagaactcat cggagggggg ctcttacctg 1680
 tggagattct gcttcggtgg gcctctagct aagctagtct atagggccaa acaggattat 1740
 aaggatcaat ttgaggatat tttgagagag tgtcctggta tttttgactc tctcaacttg 1800
 ggccatcagt ctacttttaa ccagagtatt ctgagagccc ttgacttttc tactcctggc 1860
 agaactaccg ccgcggtagc cttttttgcc tttatccttg acaaatggag tcaagaaacc 1920
 catttcagca gggattaccg tctggactgc ttagcagtag ctttgtggag aacatggagg 1980
 tgccagcgcc tgaatgcaat ctccggctac ttgccagtac agccggtaga cacgctgagg 2040
 atcctgagtc tccagtcacc ccaggaacac caacgccgcc agcagccgca gcaggagcag 2100
 cagcaagagg aggaggagga tcgagaagag aaccgagag ccggtctgga ccctccggtg 2160

gcggaggagg aggagtagct gacttgtttc ccgagctgcg ccgggtgctg actagggtctt 2220
 ccagtggacg ggagagggggg attaagcggg agaggcatga ggagactagc cacagaactg 2280
 aactgactgt cagtctgatg agccgcaggc gcccagaatc ggtgtggtgg catgagggttc 2340
 agtcgcaggg gatagatgag gtctcgggtga tgcattgagaa atattccctg gaacaagtca 2400
 agacttggtg gttggagcct gaggatgatt gggaggtagc catcaggaat tatgccaagc 2460
 tggctctgaa gccagacaag aagtacaaga ttaccaaact gattaatatc agaaattcct 2520
 gctacatttc agggaatggg gccgaggtgg agatcagtac ccaggagagg gtggccttca 2580
 gatgttgat gatgaatatg taccggggg tgggtggcat ggaggagtc acctttatga 2640
 acgcgaggtt caggggtgat ggggtataat ggggtggtctt tatggccaac accaagctga 2700
 cagtgcacgg atgtctcttc tttgggttca ataacatgtg catcgaggcc tggggcagtg 2760
 tttcagtgag gggatgcagc ttttcagcca actggatggg ggtcgtgggc agaaccaaga 2820
 gcaaggtgtc agtgaagaaa tgcctgttcg agaggtgcca cctgggggtg atgagcgagg 2880
 gcgaagccaa agtcaaacac tgcgcctcta ctgagacggg ctgctttgtg ctgatcaagg 2940
 gcaatgccca agtcaagcat aacatgatct gtggggcctc ggatgagcgc ggctaccaga 3000
 tgctgacctg cgccgggtggg aacagccata tgctggccac cgtgcatgtg acctcgcacc 3060
 cccgcaagac atggcccag ttcgagcaca acgtcatgac ccgatgcaat gtgcacctgg 3120
 ggtcccgcg aggcattgtc atgccctacc agtgcaacat gcaatttgtg aagggtgctgc 3180
 tggagcccga tgccatgtcc agagttagcc tgacgggggt gtttgacatg aatgtggagc 3240
 tgtggaaaat tctgagatat gatgaatcca agaccagggt ccgggcctgc gaatgcggag 3300
 gcaagcacgc caggcttcag cccgtgtgtg tggaggtgac ggaggacctg cgacccgatc 3360
 atttggtgtt gtcctgcaac gggacggagt tcggctccag cggggaagaa tctgactaga 3420
 gtgagtagtg tttgggggag gtggagggct tgtatgaggg gcagaatgac taaaatctgt 3480
 gtttttctgt gtgttgacgc agcatgagcg gaagcgctc ctttgagggg ggggtattca 3540
 gcccttatct gacggggcgt ctcccctctt gggcgggagt gcgtcagaat gtgatgggat 3600
 ccacggtgga cggccggccc gtgcagccc cgaactctc aaccctgacc tacgcgaccc 3660
 tgagctctc gtccgtggac gcagctgcc cgagctgctg tgcctccgcc gccagcgccg 3720
 tgcgcggaat ggccctgggc gccggctact acagctctct ggtggccaac tcgacttcca 3780
 ccaataatcc cgccagcctg aacgaggaga agctgctgct gctgatggcc cagctcgagg 3840
 ccctgacca gcgcctgggc gagctgacct agcaggtggc tcagctgcag gcggagacgc 3900
 gggccgcggt tgccacggtg aaaaccaaata aaaaaatgaa tcaataaata aacggagacg 3960
 gttgttgatt ttaacacaga gtcttgaatc tttatttgat ttttcgcgcg cggtaggccc 4020
 tggaccaccg gtctcgatca ttgagacccc ggtggatttt ttccaggacc cggtagaggt 4080
 gggcttggtg gttgaggtac atgggcatga gcccgctccc ggggtggagg tagctccatt 4140
 gcagggcctc gtgctcgggg gtggtgttgt aaatcaccca gtcattagcag gggcgaggg 4200

cgtggtgctg cacgatgtcc ttgaggagga gactgatggc cacgggcagc cccttggtgt 4260
agggtgttgac gaacctgttg agctgggagg gatgcatgcg gggggagatg agatgcatct 4320
tggcctggat cttgagattg gcgatgttcc cgcccagatc ccgccggggg ttcattgtgt 4380
gcaggaccac cagcacggtg tatccggtgc acttggggaa tttgtcatgc aacttggaag 4440
ggaaggcgtg aaagaatttg gagacgccct tgtgaccgcc caggttttcc atgcaactcat 4500
ccatgatgat ggcgatgggc ccgtgggcgg cggcctgggc aaagacgttt cgggggtcgg 4560
acacatcgta gttgtggtcc tgggtgagct cgtcataggc ctttttaatg aatttggggc 4620
ggaggggtgcc cgactggggg acgaagggtg cctcgatccc gggggcgtag ttgccctcgc 4680
agatctgcat ctcccaggcc ttgagctcgg agggggggat catgtccacc tgcggggcga 4740
tgaaaaaac ggtttccggg gcgggggaga tgagctgggc cgaaagcagg ttccggagca 4800
gctgggactt gccgcagccg gtggggccgt agatgacccc gatgaccggc tgcagggtgt 4860
agttgagggg gagacagctg ccgtcctcgc ggaggagggg ggccacctcg ttcattcatct 4920
cgcgcacatg catgttctcg cgcacgagtt ccgccaggag gcgctcgccc cccagcgaga 4980
ggagctcttg cagcgaggcg aagtttttca gcggcttgag yccgtcggcc atgggcattt 5040
tgagaggggt ctggtgcaag agttccagac ggtcccagag ctcggtgatg tgctctaggg 5100
catctcgatc cagcagacct cctcgtttcg cgggttgggg cgactgcggg agtagggcac 5160
caggcgatgg gcgtccagcg aggccagggt ccggtccttc cagggtcgca gggtcgcgt 5220
cagcgtggtc tccgtcacgg tgaaggggtg cgcgcggggc tgggcgcttg cgagggtgcg 5280
cttcaggctc atccggctgg tcgagaaccg ctcccggtcg gcgccctgcg cgtcggccag 5340
gtagcaattg agcatgagtt cgtagttgag cgcctcggcc gcgtggccct tggcgcgagg 5400
cttacctttg gaagtgtgtc cgcagacggg acagaggagg gacttgaggg cgtagagctt 5460
gggggagagg aagacggact cgggggagta ggcgtccgcg ccgcagctgg cgcagacggt 5520
ctcgcactcc acgagccagg tgaggtcggg ccggttgggg tcaaaaacga ggtttcctcc 5580
gtgctttttg atgcgtttct tacctctggt ctccatgagc tcgtgtcccc gctgggtgac 5640
aaagaggctg tccgtgtccc cgtagaccga ctttatgggc cggtcctcga gcgggggtgcc 5700
gcggtcctcg tcgtagagga accccgcca ctccgagacg aaggcccggt tccaggccag 5760
cacgaaggag gccacgtggg aggggtagcg gtcgttgtcc accagcgggt ccaccttctc 5820
cagggtatgc aagcacatgt cccctcgtc cacatccagg aagggtgattg gcttgtaagt 5880
gtagggcacg tgaccggggg tcccggccgg gggggataaa aagggggcg gcccctgctc 5940
gtcctcactg tcttccggat cgctgtccag gagcgccagc tggtggggta ggtattccct 6000
ctcgaaggct ggcataacct cggcactcag gttgtcagtt tctagaaacg aggaggattt 6060
gatattgacg gtgccgttg agacgccttt catgagcccc tcgtccatct ggtcagaaaa 6120
gacgatcttt ttgttgtcga gcttggtggc gaaggagccg tagagggcgt tggagaggag 6180
cttggcgatg gagcgcatgg tctggttctt ttccttgctg gcgcgctcct tggcggcgat 6240
gttgagctgc acgtactcgc gcgccacgca cttccattcg ggggaagacg tggtgagctc 6300

gtcgggacacg attctgaccc gccagccgcg gttgtgcagg gtgatgaggt ccacgctggt	6360
ggccacctcg ccgcgcagg gctcgttggt ccagcagagg cggccgccct tgcgcgagca	6420
gaaggggggc agcgggtcca gcatgagctc gtcggggggg tcggcgtcca cgggtgaagat	6480
gccgggcaga agctcgggggt cgaagtagct gatgcagggtg tccagatcgt ccagcgccgc	6540
ttgccagtcg cgcacggcca gcgcgcgctc gtaggggctg aggggcgtgc cccagggcagat	6600
ggggtgcgtg agcgcggagg cgtacatgcc gcagatgtcg tagacgtaga ggggctcctc	6660
gaggacccg atgtagggtg ggtagcagcg ccccccgcg atgctggcg gcacgtagtc	6720
gtacagctcg tgcgagggcg cgaggagccc cgtgccgagg ttggagcgtt gcggcttttc	6780
ggcgcggtag acgatctggc ggaagatggc gtgggagttg gaggagatgg tgggcctctg	6840
gaagatgttg aagtgggcgt ggggcaggcc gaccgagtc ctgatgaagt gggcgtagga	6900
gtcctgcagc ttggcgacga gctcggcggt gacgaggacg tccagggcg agtagtcgag	6960
ggtctcttg atgatgtcgt acttgagctg gcccttctgc ttccacagct cgcggttgag	7020
aaggaactct tcgcggtcct tccagtactc ttcgagggg aaccgcctc gatcggcacg	7080
gtaagagccc accatgtaga actggttgac ggcctttag ggcgagcagc ccttctccac	7140
ggggagggcg taagcttggt cggccttgcg cagggaggtg tgggtgaggg cgaagggtgc	7200
gcgcaccatg acctgagga actggtgctt gaagtcgagg tcgtcgcagc cgccctgctc	7260
ccagagctgg aagtccgtgc gcttcttgta ggcgggggtt ggcaaagcga aagtaacatc	7320
gttgaagagg atcttgccc cgcggggcat gaagtgcga gtgatgcgga aaggctgggg	7380
cacctcgcc cggttggtga tgacctgggc ggcgaggacg atctcgtcga agccgttgat	7440
gttgtgccc acgatgtaga gttccacgaa tcgcgggcg cccttaacgt ggggcagctt	7500
cttgagctcg tcgtagggtga gctcggcggg gtcgctgagc ccgtgctgct cgagggccca	7560
gtcggcgacg tgggggttg cgctgaggaa ggaagtccag agatccacgg ccagggcggt	7620
ctgcaagcgg tcccgggtact gacggaactg ctggccacg gccatttttt cgggggtgac	7680
gcagtagaag gtgcgggggt cgccgtgcc gcggtcccac ttgagctgga gggcgaggtc	7740
gtgggcgagc tcgacgagcg gcgggtcccc ggagagtttc atgaccagca tgaaggggac	7800
gagctgcttg ccgaaggacc ccatccaggt gtaggtttcc acatcgtagg tgaggaagag	7860
cctttcggtg cgaggatgcg agccgatggg gaagaactgg atctcctgcc accagttgga	7920
ggaatggctg ttgatgtgat ggaagtagaa atgccgacgg cgcgccgagc actcgtgctt	7980
gtgtttatac aagcgtccgc agtgctcgca acgctgcacg ggatgcacgt gctgcacgag	8040
ctgtacctg gttcctttga cgaggaatct cagtgggcag tggagcgtg gcggctgcat	8100
ctggtgctgt actacgtcct ggccatcggc gtggccatcg tctgcctcga tgggtggtcat	8160
gctgacgagc ccgcgcggga ggcagggtcca gacttcggct cggacgggtc ggagagcgag	8220
gacgagggcg cgcaggccgg agctgtccag ggtcctgaga cgctgcggag tcaggtcagt	8280
gggcagcggc ggcgcgcggt tgacttgag gagcttttcc agggcgcgcg ggagggtccag	8340

atggtacttg	atctccacgg	cgccgttggt	ggcgacgtcc	acggcttgca	gggtcccgtg	8400
cccctggggc	gccaccaccg	tgccccgttt	cttcttgggc	gctgcttcca	tgccggtcag	8460
aagcggcggc	gaggacgcgc	gccgggcggc	aggggcggct	cgggacccgg	aggcaggggc	8520
ggcaggggca	cgtcggcgcc	gcgcgcgggc	aggttcttgt	actgcgcccg	gagaagactg	8580
gcgtgagcga	cgacgcgacg	gttgacgtcc	tggatctgac	gcctctgggt	gaaggccacg	8640
ggacccgtga	gtttgaacct	gaaagagagt	tcgacagaat	caatctcggg	atcgttgacg	8700
gcggcctgcc	gcaggatctc	ttgcacgtcg	cccgagttgt	cctggtaggc	gatctcggtc	8760
atgaactgct	cgatctcctc	ctcctgaagg	tctccgcggc	cggcgcgctc	gacggtggcc	8820
gcgaggtcgt	tggagatgcg	gcccattgac	tgcgagaagg	cgttcatgcc	ggcctcgttc	8880
cagacgcggc	tgtagaccac	ggctccgtcg	gggtcgcgcg	cgcgcatgac	cacctgggcg	8940
aggttgagct	cgacgtggcg	cgtgaagacc	gcgtagttgc	agaggcgctg	gtagaggtag	9000
ttgagcgtgg	tggcgatgtg	ctcgggtgacg	aagaagtaca	tgatccagcg	gcggagcggc	9060
atctcgctga	cgtcgcccag	ggcttccaag	cgctccatgg	cctcgtagaa	gtccacggcg	9120
aagttgaaaa	actgggagtt	gcgcgcggag	acggtcaact	cctcctccag	aagacggatg	9180
agctcagcga	tggtggcgcg	cacctcgcg	tcgaaggccc	cggggggctc	ctcttcttcc	9240
atctcttcct	cctccactaa	catctcttct	acttcctcct	caggaggcgg	cggcggggga	9300
ggggccctgc	gtcgcggcg	gcgcacgggc	agacggtcga	tgaagcgctc	gatggtctcc	9360
ccgcgcgggc	gacgcattgt	ctcgggtgacg	gcgcgcccgt	cctcgcgggg	ccgcagcgtg	9420
aagacgccgc	cgcgcatctc	cagggtggccg	ccgggggggt	ctccgttggg	cagggagagg	9480
gcgctgacga	tgcattttat	caattggccc	gtagggactc	cgcgcaagga	cctgagcgtc	9540
tcgagatcca	cgggatccga	aaaccgctga	acgaaggctt	cgagccagtc	gcagtcgcaa	9600
ggtaggctga	gcccggtttc	ttgttcttcg	gggatttcgg	gaggcgggcg	ggcgatgctg	9660
ctggtgatga	agttgaagta	ggcggctcctg	agacggcgga	tggtggcgag	gagcaccagg	9720
tccttgggcc	cggcttgctg	gatgcgcaga	cggtcggcca	tgccccaggc	gtggtcctga	9780
cacctggcga	ggctcttgta	gtagtcctgc	atgagccgct	ccacgggcac	ctcctcctcg	9840
cccgcgcggc	cgtgcatgcg	cgtgagcccg	aaccgcgct	ggggctggac	gagcgccagg	9900
tcggcgacga	cgcgctcggc	gaggatggcc	tgctgtatct	gggtgagggt	ggtctggaag	9960
tcgtcgaagt	cgacgaagcg	gtggtaggct	ccggtgttga	tggtatagga	gcagttggcc	10020
atgacggacc	agttgacggg	ctggtggccg	ggtcgcacga	gctcgtggta	cttgaggcgc	10080
gagtaggcgc	gcgtgtcgaa	gatgtagtcg	ttgcaggtgc	gcacgaggta	ctggtatccg	10140
acgaggaagt	gcggcggcg	ctggcggtag	agcggccatc	gctcgggtggc	gggggcggcg	10200
ggcgcgaggt	cctcgagcat	gaggcggtgg	tagccgtaga	tgtacctgga	catccagggtg	10260
atgccggcgg	cgggtgggga	ggcgcgcggg	aactcgcgga	cgcggttcca	gatgttgcg	10320
agcggcagga	agtagttcat	ggtggcccg	gtctggcccc	tgaggcgcg	gcagtcgtgg	10380
atgctctaga	catacgggca	aaaacgaaag	cggtcagcgg	ctcgactccg	tggcctggag	10440

gctaagcgaa	cggggttgggc	tgcgcgtgta	ccccggttcg	aatctcgaat	caggctggag	10500
ccgcagctaa	cgtggtactg	gcactcccgt	ctcgacccaa	gcctgctaac	gaaacctcca	10560
ggatacggag	gcgggtcgtt	ttttggcctt	ggtcgcgtgt	catgaaaaac	tagtaagcgc	10620
ggaaagcgac	cgcccgcgat	ggctcgcctg	cgtagtctgg	agaaagaatc	gccagggttg	10680
cgttgcggtg	tgccccggtt	cgagcctcag	cgctcggcgc	cggccggatt	ccgcggctaa	10740
cgtgggcgtg	gctgccccgt	cgtttccaag	acccttagc	cagccgactt	ctccagttac	10800
ggagcgagcc	cctctttttc	ttgtgttttt	gccagatgca	tcccgtactg	cggcagatgc	10860
gccccaccc	tccacctcaa	ccgcccctac	cgccgcagca	gcagcaacag	ccggcgcttc	10920
tgccccgcc	ccagcagcag	ccagccacta	ccgcggcggc	cgccgtgagc	ggagccggcg	10980
ttcagtatga	cctggccttg	gaagagggcg	aggggctggc	gcggctgggg	gcgtcgtcgc	11040
cggagcggca	cccgcgcgtg	cagatgaaaa	gggacgctcg	cgaggcctac	gtgccaagc	11100
agaacctgtt	cagagacagg	agcggcgagg	agcccgagga	gatgcgcgcc	tcccgttcc	11160
acgcggggcg	ggagctgcgg	cgcggcctgg	accgaaagcg	ggtgctgagg	gacgaggatt	11220
tcgaggcgga	cgagctgacg	gggatcagcc	ccgcgcgcgc	gcacgtggcc	gcggccaacc	11280
tggtcacggc	gtacgagcag	accgtgaagg	aggagagcaa	cttccaaaaa	tccttcaaca	11340
accacgtgcg	cacgctgac	gcgcgcgagg	aggtgaccct	gggcctgatg	cacctgtggg	11400
acctgctgga	ggccatcgtg	cagaacccca	cgagcaagcc	gctgacggcg	cagctgtttc	11460
tggttggtgca	gcacagtcgg	gacaacgaga	cgttcaggga	ggcgtgctg	aatatcaccg	11520
agcccgaggg	ccgctggctc	ctggacctgg	tgaacattct	gcagagcatc	gtggtgcagg	11580
agcgcgggct	gccgctgtcc	gagaagctgg	cggctatcaa	cttctcggtg	ctgagcctgg	11640
gcaagtacta	cgctaggaag	atctacaaga	ccccgtacgt	gcccatagac	aaggagggtga	11700
agatcgacgg	gttttacatg	cgcatgacct	tgaaagtgct	gacctgagc	gacgatctgg	11760
gggtgtaccg	caacgacagg	atgcaccgcg	cggtagcgc	cagccgccgg	cgcgagctga	11820
gcgaccagga	gctgatgcac	agcctgcagc	gggccctgac	cggggccggg	accgaggggg	11880
agagctactt	tgacatgggc	gcggacctgc	gctggcagcc	cagccgccgg	gccttggaag	11940
ctgccggcgg	ttccccctac	gtggaggagg	tggacgatga	ggaggaggag	ggcgagtacc	12000
tggaagactg	atggcgcgac	cgtatttttg	ctagatgcag	caacagccac	cgctcctga	12060
tcccgcgatg	cgggcggcgc	tgcagagcca	gccgtccggc	attaactcct	cggacgattg	12120
gaccagggcc	atgcaacgca	tcatggcgct	gacgaccgc	aatcccgaag	cctttagaca	12180
gcagcctcag	gccaaaccgg	tctcggccat	cctggaggcc	gtggtgccct	cgcgctcgaa	12240
ccccacgcac	gagaaggtgc	tggccatcgt	gaacgcgctg	gtggagaaca	aggccatccg	12300
cggcgacgag	gccgggctgg	tgtacaacgc	gctgctggag	cgcgtggccc	gctacaacag	12360
caccaacgtg	cagacgaacc	tggaccgcat	ggtgaccgac	gtgcgcgagg	cgggtgtcgca	12420
gcgcgagcgg	ttccaccgcg	agtcgaacct	gggctccatg	gtggcgctga	acgccttcct	12480

gagcacgcag cccgccaacg tgccccgggg ccaggaggac tacaccaact tcatcagcgc 12540
gctgcggctg atggtggccg aggtgcccc aagcgcaggtg taccagtcgg ggccggacta 12600
cttcttccag accagtcgcc agggcttgca gaccgtgaac ctgagccagg ctttcaagaa 12660
cttgaggga ctgtggggcg tgcaggcccc ggtcggggac cgcgcgacgg tgtcagacct 12720
gctgacgccg aactcgcgcc tgctgctgct gctggtggcg cccttcacgg acagcggcag 12780
cgtgagccgc gactcgtacc tgggctacct gcttaacctg taccgcgagg ccatcgggca 12840
ggcgcacgtg gacgagcaga cctaccagga gatcacccac gtgagccgcg cgctgggcca 12900
ggaggacccg ggcaacctgg aggccaccct gaacttcctg ctgaccaacc ggtcgcagaa 12960
gatcccgccc cagtacgcgc tgagcaccga ggaggagcgc atcctgcgct acgtgcagca 13020
gagcgtgggg ctgttcctga tgcaggaggg ggccacgccc agcgcgcgc tcgacatgac 13080
cgcgcgcaac atggagccca gcatgtacgc tcgcaaccgc ccgttcatca ataagctgat 13140
ggactacttg catcgggchg ccgccatgaa ctcggactac tttaccaacg ccatcttgaa 13200
cccgcactgg ctcccgcgc cggggttcta cacgggchgag tacgacatgc ccgaccccaa 13260
cgacgggttc ctgtgggacg acgtggacag cagcgtgttc tcgccgcgc ccgccaccac 13320
cgtgtggaag aaagagggcg gggaccggcg gccgtcctcg gcgctgtccg gtcgcgcggg 13380
tgctgcccg gcggtgcctg aggccgccag ccccttccc agcctgccct tttcgtgaa 13440
cagcgtgcgc agcagcgagc tgggtcggct gacgcggccg cgcctgctgg gcgaggagga 13500
gtacctgaac gactccttgt tgaggcccga gcgcgagaag aacttcccca ataacgggat 13560
agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgcacgagc acagggacga 13620
gccccgagct agcagcgagc caggcacccg tagacgccag cgacacgaca ggcagcgggg 13680
tctggtgtgg gacgatgagg attccgccga cgacagcagc gtgttggact tgggtgggag 13740
tggtggtggt aaccggttcg ctcaattgcg ccccggtatc gggcgccctga tgtaagaatc 13800
tgaaaaaata aaaaacggta ctcaccaagg ccatggcgac cagcgtgcgt tcttctctgt 13860
tgtttgtagt agtatgatga ggcgcgtgta cccggagggt cctcctccct cgtacgagag 13920
cgtgatgcag caggcggtgg cggcgcgat gcagcccccg ctggaggcg cttacgtgcc 13980
cccgcggtac ctggcgcta cggagggcg gaacagcatt cgttactcgg agctggcacc 14040
cttgtagcat accaccggt tgtacctggt ggacaacaag tcggcgga tgcctcgt 14100
gaactaccag aacgaccaca gcaacttcct gaccaccgtg gtgcagaaca acgatttcac 14160
ccccacggag gccagaccc agaccatcaa ctttgacgag cgctcgcggt gggcgggcca 14220
gctgaaaacc atcatgcaca ccaacatgcc caacgtgaac gagttcatgt acagcaacaa 14280
gttcaaggcg cgggtgatgg tctcgcgcaa gaccccaat ggggtcgcgg tggatgagaa 14340
ttatgatggt agtcaggacg agctgactta cgagtgggtg gagtttgagc tgcccgagg 14400
caacttctcg gtgacatga ccatcgatct gatgaacaac gccatcatcg acaactactt 14460
ggcgggtggg cgtcagaacg ggggtgctgga gagcgacatc ggcgtgaagt tcgacacgcg 14520
caacttccgg ctgggctggg acccgtgac cgagctggtg atgccggcg tgtacaccaa 14580

cgaggccttc caccgagaca tcgtcctgct gcccggtgct ggcgtggact tcaccgagag 14640
 ccgcctcagc aacctgctgg gcatccgcaa gcggcagccc ttccaggagg gcttccagat 14700
 cctgtacgag gacctggagg ggggcaacat ccccgcgctc ttggatgtcg aagcctatga 14760
 gaaaagcaag gaggaggccg ccgcagcggc gaccgcagcc gtggccaccg cctctaccga 14820
 ggtgcggggc gataattttg ctagcgccgc ggcagtggcc gaggcggctg aaaccgaaag 14880
 taagatagtc atccagccgg tggagaagga cagcaaggac aggagctaca acgtgctcgc 14940
 ggacaagaaa aacaccgcct accgcagctg gtacctggcc tacaactacg gcgacccccga 15000
 gaagggcgtg cgctcctgga cgctgctcac cacctcggac gtcacctgcg gcgtggagca 15060
 agtctactgg tcgctgcccg acatgatgca agaccgggtc accttcgct ccacgcgtca 15120
 agttagcaac taccgggtgg tggcgccga gctcctgccc gtctactcca agagcttctt 15180
 caacgagcag gccgtctact cgcagcagct gcgcgccttc acctcgctca cgcacgtctt 15240
 caaccgcttc cccgagaacc agatcctcgt ccgcccgcgc gcgccacca ttaccaccgt 15300
 cagtgaaaac gttcctgctc tcacagatca cgggaccctg ccgctgcgca gcagtatccg 15360
 gggagtccag cgcgtgaccg tctactgacgc cagacgccgc acctgcccct acgtctacaa 15420
 ggccctgggc gtagtcgcgc cgcgcgtcct ctcgagccgc accttctaaa aaatgtccat 15480
 tctcatctcg cccagtaata acaccggttg gggcctgcgc gcgccagca agatgtacgg 15540
 aggcgctcgc caacgctcca cgcaacaccc cgtgcgcgtg cgcgggcact tccgcgctcc 15600
 ctggggcgcc ctcaagggcc gcgtgcgctc gcgcaccacc gtcgacgacg tgatcgacca 15660
 ggtggtggcc gacgcgcgca actacacgcc cgccgccgcg cccgcctcca ccgtggacgc 15720
 cgtcatcgac agcgtggtgg ccgatgcgcg ccggtacgcc cgcgccaaga gccggcgggc 15780
 gcgcatcgcc cggcggcacc ggagcacccc cgccatgcgc gcggcgcgag ccttgctgcg 15840
 cagggccagg cgcacgggac gcagggccat gctcagggcg gccagacgcg cggcctccgg 15900
 cagcagcagc gccggcagga cccgcagacg cgcggccacg gcggcgggcg cggccatcgc 15960
 cagcatgtcc cgcccgcggc gcggcaacgt gtactgggtg cgcgacgccg ccaccggtgt 16020
 gcgcgtgccc gtgcgcaccc gccccctcg cacttgaaga tgctgacttc gcgatgttga 16080
 tgtgtcccag cggcgaggag gatgtccaag cgaaaataca aggaagagat gctccagggtc 16140
 atcgcgcctg agatctacgg ccccgcggtg aaggaggaaa gaaagccccg caaactgaag 16200
 cgggtcaaaa aggacaaaaa ggaggaggaa gatgtggacg gactggtgga gtttgtgcgc 16260
 gagttcggcc cccggcgggc cgtgcagtgg cgcgggcgga aagtgaacc ggtgctgcgg 16320
 cccggcacca cgggtggtctt cacgcccggc gagcgttccg gctccgcctc caagcgtcc 16380
 tacgacgagg tgtacgggga cgaggacatc ctcgagcagg cggtcgagcg tctgggagcg 16440
 tttgcttacg gcaagcgcag ccgccccgcg cccttgaaag aggaggcggt gtccatccccg 16500
 ctggaccacg gcaacccac gccgagcctg aagccggtga ccctgcagca ggtgctgccg 16560
 agcgcggcgc cgcgcggggg cttcaagcgc gagggcgggc aggatctgta cccgaccatg 16620

cagctgatgg	tgcccaagcg	ccagaagctg	gaggacgtgc	tggagcacat	gaaggtggac	16680
cccgaggtgc	agcccgaggt	caaggtgctg	cccatcaagc	aggtggcccc	gggcctgggc	16740
gtgcagaccg	tggacatcaa	gatccccacg	gagcccatgg	aaacgcagac	cgagcccgtg	16800
aagcccagca	ccagcaccat	ggaggtgcag	acggatccct	ggatgccggc	gccggcttcc	16860
accactcgcc	gaagacgcaa	gtacggcgcg	gccagcctgc	tgatgcccac	ctacgcgctg	16920
catccttcca	tcatccccac	gccgggctac	cgcggcacgc	gcttctaccg	cggctacacc	16980
agcagccgcc	gcaagaccac	cacccgcccgc	cgccgtcgtc	gcacccgccg	cagcagcacc	17040
gcgacttccg	ccgccgccct	ggtgcggaga	gtgtaccgca	gcgggcgcga	gcctctgacc	17100
ctgccgcgcg	cgcgctacca	cccagagcatc	gccatttaac	tctgccgtcg	cctcctactt	17160
gcagatatgg	ccctcacatg	ccgcctccgc	gtccccatta	cgggctaccg	aggaagaaaag	17220
ccgcgccgta	gaaggctgac	ggggaacggg	ctgcgtcgcc	atcaccaccg	gcggcgggcg	17280
gccatcagca	agcggttggg	gggaggcttc	ctgcccgcgc	tgatccccat	catcgccgcg	17340
gcgatcgggg	cgatccccgg	catagcttcc	gtggcggtgc	aggcctctca	gcgccactga	17400
gacacagctt	ggaaaatttg	taataaaaaa	atggactgac	gctcctgggtc	ctgtgatgtg	17460
tgtttttaga	tggaagacat	caatttttcg	tccctggcac	cgcgacacgg	cacgcggccg	17520
tttatgggca	cctggagcga	catcggaac	agccaactga	acggggggcg	cttcaattgg	17580
agcagtctct	ggagcgggct	taagaatttc	gggtccacgc	tcaaaaccta	tggcaacaag	17640
gcgtggaaca	gcagcacagg	gcaggcgctg	agggaacagc	tgaaagagca	gaacttccag	17700
cagaaggtgg	tcgatggcct	ggcctcgggc	atcaacgggg	tggaggacct	ggccaaccag	17760
gccgtgcaga	aacagatcaa	cagccgcctg	gacgcggtcc	cgcccgcggg	gtccgtggag	17820
atgccccagg	tggaggagga	gctgcctccc	ctggacaagc	gcggcgacaa	gcgaccgcgt	17880
cccgacgcgg	aggagacgct	gctgacgcac	acggacgagc	cgcccccgta	cgaggaggcg	17940
gtgaaactgg	gtctgcccac	cacgcggccc	gtggcgccctc	tggccaccgg	ggtgctgaaa	18000
cccagcagca	gcagccagcc	cgcgaccctg	gacttgccctc	cgcttgcttc	ccgcccctcc	18060
acagtggcta	agcccctgcc	gccgggtggc	gtcgcgtcgc	gcgccccccg	aggccgcccc	18120
caggcgaact	ggcagagcac	tctgaacagc	atcgtgggtc	tgggagtgca	gagtgtgaag	18180
cgccgcccgt	gctattaaaa	gacactgtag	cgcttaactt	gcttgtctgt	gtgtatatgt	18240
atgtccgccg	accagaagga	ggaagaggcg	cgtcgccgag	ttgcaagatg	gccaccccat	18300
cgatgctgcc	ccagtgggcg	tacatgcaca	tcgccggaca	ggacgcttcg	gagtacctga	18360
gtccgggtct	ggtgcagttc	gcccgcgcca	cagacaccta	cttcagtctg	gggaacaagt	18420
ttaggaaccc	cacggtggcg	cccacgcacg	atgtgaccac	cgaccgcagc	cagcggctga	18480
cgctgcgctt	cgtagccgctg	gaccgcgagg	acaacaccta	ctcgtaaaa	gtgcgctaca	18540
cgctggccgt	gggcgacaac	cgcgtagctg	acatggccag	cacctacttt	gacatccgcg	18600
gcgtgctgga	tcggggggcc	agcttcaaac	cctactccgg	caccgcctac	aacagcctgg	18660
ctcccaaggg	agcgcccaac	acttgccagt	ggacatataa	agctgggtgat	actgatacag	18720

aaaaaaccta tacatatgga aatgcacctg tgcaaggcat tagcattaca aaggatggta 18780
ttcaacttgg aactgacagc gatggtcagg caatctatgc agacgaaact tatcaaccag 18840
agcctcaagt gggatgatgct gaatggcatg acatcactgg tactgatgaa aaatatggag 18900
gcagagctct taagcctgac accaaaatga agccttgcta tggttctttt gccaaagccta 18960
ccaataaaga aggaggccag gcaaattgtga aaaccgaaac aggcgggtacc aaagaatatg 19020
acattgacat ggcatttctt gataatcgaa gtgcagctgc cgccggccta gccccagaaa 19080
ttgttttgta tactgagaat gtggatctgg aaactccaga taccatatt gtatacaagg 19140
caggtacaga tgacagtagc tcttctatca atttgggtca gcagtccatg cccaacagac 19200
ccaactacat tggcttcaga gacaacttta tcggtctgat gtactacaac agcactggca 19260
atatgggtgt actggctgga caggcctccc agctgaatgc tgtggtggac ttgcaggaca 19320
gaaacaccga actgtcctac cagctcttgc ttgactctct gggtagacaga accagggtatt 19380
tcagtatgtg gaatcaggcg gtggacagtt atgaccccga tgtgctgatt attgaaaatc 19440
acggtgtgga ggatgaactt cctaactatt gcttccccct ggatgctgtg ggtagaactg 19500
atacttacca gggaattaag gccaatggtg ataatacaac cacctggacc aaagatgata 19560
ctgttaatga tgctaataa ttgggcaagg gcaatcctt cgccatggag atcaacatcc 19620
aggccaacct gtggcggaac ttcctctacg cgaacgtggc gctgtacctg cccgactcct 19680
acaagtacac gccggccaac atcacgctgc ccaccaacac caacacctac gattacatga 19740
acggccgcgt ggtggcgccc tcgctggtgg acgcctacat caacatcggg gcgcgctggt 19800
cgctggaccc catggacaac gtcaaccctt tcaaccacca ccgcaacgcg ggctgctgat 19860
accgctccat gctcctgggc aacggcgct acgtgccctt ccacatccag gtgccccaaa 19920
agtttttcgc catcaagagc ctctgctcc tgcccggtc ctacacctac gagtggaact 19980
tccgcaagga cgtcaacatg atcctgcaga gtcctctcg caacgacctg cgcacggacg 20040
gggctccat cgccttcacc agcatcaacc tctacgccac cttcttcccc atggcgacac 20100
acaccgcctc cagctcgag gccatgctgc gcaacgacac caacgaccag tccttcaacg 20160
actacctctc ggcggccaac atgctctacc ccacccggc caacgccacc aacgtgcca 20220
tctccatccc ctgcgcaac tgggcccct tccgcggtg gtccttcacg cgcctcaaga 20280
cccgcgagac gccctcgtc ggctccgggt tcgacccta cttcgtctac tcgggtcca 20340
tcccctacct cgacggcacc ttctacctca accacacctt caagaagggt tccatcacct 20400
tcgactctc cgtcagctgg ccgggaacg accgcctcct gacgccaac gagttcgaaa 20460
tcaagcgac cgtcgacgga gagggtgaca acgtggcca gtgcaacatg accaaggact 20520
ggttcctggt ccagatgctg gccactaca acatcggtc ccagggttc tacgtgcccg 20580
agggtacaa ggaccgatg tactccttct tccgcaactt ccagcccatg agccgccagg 20640
tcgtggacga ggtcaactac aaggactacc aggcgtcac cctggcctac cagcacaaca 20700
actcgggctt cgtcggctac ctgcgccc caatgcgcca gggccagccc taccgcgcca 20760

actacccta cccgctcatc ggcaagagcg ccgtcgccag cgtcaccag aaaaagtcc 20820
 tctgcgaccg ggtcatgtgg cgcaccccct tctccagcaa cttcatgtcc atgggagcgc 20880
 tcaccgacct cggccagaac atgctctacg ccaactccgc ccacgcgcta gacatgaatt 20940
 tcgaagtcca ccccatggat gagtcacccc ttctctatgt tgtcttcgaa gtcttcgacg 21000
 tcgtccgagt gcaccagccc caccgcggcg tcatcgaggc cgtctacctg cgcacgccct 21060
 tctcggccgg caacgccacc acctaacct cttgcttctt gcaagatgac ggcctgcgcg 21120
 ggctccggcg agcaggagct cagggccatc ctccgcgacc tgggctgcgg gccctgcttc 21180
 ctgggcacct tcgacaagcg cttcccggga ttcattggccc cgcacaagct ggcctgcgcc 21240
 atcgtcaaca cggccggccg cgagaccggg ggcgagcact ggctggcctt cgcctggaac 21300
 ccgcgctccc acacctgcta cctcttcgac cccttcgggt tctcggacga gcgcctcaag 21360
 cagatctacc agttcgagta cgagggcctg ctgcgtcgca gcgccctggc caccgaggac 21420
 cgctgcgtca ccctggaaaa gtccaccag accgtgcagg gtccgcgctc ggccgcctgc 21480
 gggctcttct gctgcatgtt cctgcacgcc ttcgtgcact ggcccgaccg ccccatggac 21540
 aagaaccca ccatgaactt gctgacggg gtgccaacg gcatgctcca gtcgccccag 21600
 gtggaaccca ccctgcgccg caaccaggag gcgctctacc gcttcctcaa cgcctactcc 21660
 gcctactttc gctcccaccg cgcgcgcacg gagaaggcca ccgccttcga ccgcatgaat 21720
 caagacatgt aatccggtgt gtgtatgtga atgctttatt catcataata aacagcacat 21780
 gtttatgcc cttctctga ggctctgact ttatttagaa atcgaagggg ttctgcccgc 21840
 tctcggcatg gcccgcgggc agggatacgt tgcggaactg gtacttgggc agccacttga 21900
 actcggggat cagcagcttc ggcacgggga ggtcggggaa cgagtcgctc cacagcttgc 21960
 gcgtgagttg cagggcgccc agcaggctcg gcgcggagat cttgaaatcg cagttgggac 22020
 ccgcgttctg cgcgcgagag ttacggtaca cggggttgca gcactggaac accatcaggg 22080
 ccgggtgctt cacgctcgcc agcaccgtcg cgtcgggtgat gccctccacg tccagatcct 22140
 cggcgttggc catcccgaag ggggtcatct tgcaggctctg ccgccccatg ctgggcacgc 22200
 agccgggctt gtggttgcaa tcgagtgca gggggatcag catcatctgg gcctgctcg 22260
 agctcatgcc cgggtacatg gccttcata aagcctccag ctggcggaag gcctgctgcg 22320
 ccttgccgcc ctcgggtgaag aagacccgc agacttgct agagaactgg ttggtggcgc 22380
 agccagcgtc gtgcacgcag cagcgcgcgt cgttggtggc cagctgcacc acgctgcgcc 22440
 cccagcgggt ctgggtgatc ttggcccggt cggggttctc cttcagcgcg cgtgcccgt 22500
 tctcgtcgc cacatccatc tcgacgtgt gtccttctg gatcatcacg gtcccgtgca 22560
 ggcaccgcag cttgccctcg gcctcgggtc acccgtgcag ccacagcgcg cagccggtgc 22620
 tctcccagtt cttgtgggcg atctgggagt gcgagtgac gaagccctgc aggaagcggc 22680
 ccatcatcgt ggtcagggtc ttgttgctgg tgaaggctcag cggaatgccg cgggtgctcct 22740
 cgttcacata caggtggcag atacggcggt acacctgcc ctgctcgggc atcagctgga 22800
 aggcggactt caggtcgctc tccacgcggt accggtccat cagcagcgtc atcacttcca 22860

tgcccttctc	ccaggccgaa	acgatcggca	ggctcagggg	gttcttcacc	gttgatcatct	22920
tagtcgccgc	cgccgaagtc	aggggggtcgt	tctcgtccag	ggtctcaaac	actcgtttgc	22980
cgctcttctc	ggtgatgcgc	acgggggggaa	agctgaagcc	cacggccgcc	agctcctcct	23040
cggcctgcct	ttcgtcctcg	ctgtcctggc	tgatgtcttg	caaaggcaca	tgcttggtct	23100
tgcgggggttt	ctttttgggc	ggcagagggc	gcggcggaga	cgtgctgggc	gagcgcgagt	23160
tctcgtcac	cacgactatt	tcttctcctt	ggccgtcgtc	cgagaccacg	cggcggtagg	23220
catgcctctt	ctggggcaga	ggcggagggc	acgggctctc	gcggttcggc	gggcggctgg	23280
cagagcccct	tccgcgttcg	ggggtgcgct	cctggcggcg	ctgctctgac	tgacttcctc	23340
cgcggccggc	cattgtgttc	tcctagggag	caagcatgga	gactcagcca	tcgtcgccaa	23400
catcgccatc	tgcccccgcc	gccgccgacg	agaaccagca	gcagcagaat	gaaagcttaa	23460
ccgccccgcc	gcccagcccc	acctccgacg	ccgcagcccc	agacatgcaa	gagatggagg	23520
aatccatcga	gattgacctg	ggctacgtga	cgcccgcgga	gcacgaggag	gagctggcag	23580
cgcgcttttc	agccccggaa	gagaaccacc	aagagcagcc	agagcaggaa	gcagagagcg	23640
agcagaacca	ggctgggctc	gagcatggcg	actacctgag	cggggcagag	gacgtgctca	23700
tcaagcatct	ggcccgccaa	tgcatcatcg	tcaaggacgc	gctgctcgac	cgcgccgagg	23760
tgcccctcag	cgtggcggag	ctcagccgcg	cctacgagcg	caacctcttc	tcgccgcgcg	23820
tgccccccaa	gcgccagccc	aacggcacct	gcgagcccaa	cccgcgcctc	aacttctacc	23880
cggctttcgc	ggtgcccgag	gccctggcca	cctaccacct	ctttttcaag	aaccaaagga	23940
tccccgtctc	ctgccgcgcc	aaccgcaccc	gcgccgacgc	cctgctcaac	ctgggccccg	24000
gcgccccgcct	acctgatatc	gcctccttgg	aagaggttcc	caagatcttc	gagggtcttg	24060
gcagcgacga	gactcggggc	gcgaacgctc	tgcaaggaag	cggagaggag	catgagcacc	24120
acagcgccct	ggtggagttg	gaaggcgaca	acgcgcgcct	ggcggtcctc	aagcgcacgg	24180
tcgagctgac	ccacttcgcc	tacccggcgc	tcaacctgcc	ccccaaaggtc	atgagcgccg	24240
tcatggacca	ggtgctcatc	aagcgcgcct	cgcccctctc	ggaggaggag	atgcaggacc	24300
ccgagagctc	ggacgagggc	aagcccgtgg	tcagcgacga	gcagctggcg	cgctggctgg	24360
gagcgagtag	cacccccag	agcctggaag	agcggcgcaa	gctcatgatg	gccgtggtcc	24420
tggtgaccgt	ggagctggag	tgtctgcgcc	gcttcttcgc	cgacgcggag	accctgcgca	24480
aggtcgagga	gaacctgcac	tacctttca	gacacgggtt	cgtgcgccag	gcctgcaaga	24540
tctccaacgt	ggagctgacc	aacctgggtt	cctacatggg	catcctgcac	gagaaccgcc	24600
tggggcagaa	cgtgctgcac	accaccctgc	gcggggaggc	ccgccgcgac	tacatccgcg	24660
actgcgtcta	cctgtacctc	tgccacacct	ggcagacggg	catgggctg	tggcagcagt	24720
gcctggagga	gcagaacctg	aaagagctct	gcaagctcct	gcagaagaac	ctcaaggccc	24780
tgtggaccgg	gttcgacgag	cgcaccaccg	ccgcggacct	ggccgacctc	atcttccccg	24840
agcgcctgcg	gctgacgctg	cgcaacgggc	tgcccgactt	tatgagccaa	agcatgttgc	24900

aaaacttttcg ctctttcatc ctcgaaacgct ccgggatcct gcccgccacc tgctccgcgc 24960
tgccctcgga cttcgtgccg ctgaccttcc gcgagtgccccc cccgccgctc tggagccact 25020
gctacctgct gcgcctggcc aactacctgg cctaccactc ggacgtgatc gaggacgtca 25080
gcggcgaggg cctgctcgag tgccactgcc gctgcaacct ctgcacgccg caccgctccc 25140
tggcctgcaa cccccagctg ctgagcgaga cccagatcat cggcaccttc gagttgcaag 25200
gccccggcga gggcaagggg ggtctgaaac tcaccccggg gctgtggacc tcggcctact 25260
tgcgcaagtt cgtgcccag gactaccatc ctttcgagat caggttctac gaggaccaat 25320
cccagccgcc caaggccag ctgtcggcct gcgtcatcac ccagggggcc atcctggccc 25380
aattgcaagc catccagaaa tcccgcgaag aatttctgct gaaaaagggc cacgggggtct 25440
acttggaacc ccagaccgga gaggagctca accccagctt ccccgaggat gccccgagga 25500
agcagcaaga agctgaaagt ggagctgccg ccgccgccgg aggatttgga ggaagactgg 25560
gagagcagtc aggcagagga ggaggagatg gaagactggg acagcactca ggcagaggag 25620
gacagcctgc aagacagtct ggaggaggaa gacgaggtgg aggaggcaga ggaagaagca 25680
gccgccgcca gaccgtcgtc ctcggcgag gaggagaaag caagcagcac ggataccatc 25740
tccgctccgg gtcgggggtc cggcgggccg gccacagta gatgggacga gaccgggcgc 25800
ttcccgaacc ccaccacca gaccggtaag aaggagcggc agggatacaa gtcctggcgg 25860
gggcacaaaa acgccatcgt ctctgcttg caagcctgcg ggggcaacat ctccttcacc 25920
cggcgctacc tgctcttcca ccgcggggtg aacttcccc gcaacatctt gcattactac 25980
cgtcacctcc acagccccta ctactgtttc caagaagagg cagaaacca gcagcagcag 26040
cagcagcaga aaaccagcgg cagcagctag aaaatccaca gcggcggcag gtggactgag 26100
gatcgcgcg aacgagccgg cgcagaccgg ggagctgagg aaccggatct tccccaccct 26160
ctatgccatc ttccagcaga gtcgggggca agagcaggaa ctgaaagtca agaaccgttc 26220
tctgcgctcg ctcacccgca gttgtctgta tcacaagagc gaagaccaac ttcagcgcac 26280
tctcgaggac gccgaggctc tcttcaacaa gtactgcgcg ctcactctta aagagtagcc 26340
cgcgcccgcc cacacacgga aaaaggcggg aattacgtca ccacctgcgc cttcgccccg 26400
accatcatca tgagcaaaga gattccacg cttacatgt ggagctacca gccccagatg 26460
ggcctggccg ccggcgccgc ccaggactac tccaccgcga tgaactggct cagtgccggg 26520
cccgcgatga tctcacgggt gaatgacatc cgcgccacc gaaaccagat actcctagaa 26580
cagtcagcga tcaccgccac gccccgccat caccttaatc cgcgtaattg gcccgcgcc 26640
ctgggtgtacc aggaaattcc ccagcccacg accgtactac ttccgcgaga cggccaggcc 26700
gaagtccagc tgactaactc aggtgtccag ctggccggcg gcgccgccct gtgtcgtcac 26760
cgccccgctc agggataaaa gcggctggtg atccgaggca gaggcacaca gctcaacgac 26820
gaggtggtga gctcttcgct gggctctgca cctgacggag tcttccaact cgccggatcg 26880
gggagatctt cttcacgcc tcgtcaggcc gtcctgactt tggagagttc gtcctcgag 26940
ccccgctcgg gtggcatcgg cactctccag ttcgtggagg agttcactcc ctcggtctac 27000

ttcaaccctt tctccggctc ccccggccac taccgggacg agttcatccc gaacttcgac 27060
gccatcagcg agtcgggtgga cggctacgat tgaatgtccc atgggtggcgc ggctgacctt 27120
gctcggcttc gacacctgga ccactgccgc cgcttccgct gcttcgctcg ggatctcgcc 27180
gagtttgctt actttgagct gcccggaggag caccctcagg gccgggcca cggagtgcgg 27240
atcgctcgtc aaggggggtct cgactccac ctgcttcgga tcttcagcca gcgtccgatc 27300
ctggccgagc gcgagcaagg acagaccctt ctgaccctgt actgcatctg caaccacccc 27360
ggcctgcatg aaagtctttg ttgtctgctg tgtactgagt ataataaaag ctgagatcag 27420
cgactactcc ggacttccgt gtgttcctgc tatcaaccag tccctgttct tcaccgggaa 27480
cgagaccgag ctccagctcc agtgtaagcc ccacaagaag tacctcacct ggctgtttcca 27540
gggctctccg atcgccgttg tcaaccactg cgacaacgac ggagtcctgc tgagcggccc 27600
tgccaacctt actttttcca cccgcagaag caagctccag ctcttccaac ccttcctccc 27660
cgggacctat cagtgcgtct cgggacctg ccatcacacc ttccacctga tcccgaatac 27720
cacagcgtcg ctccccgcta ctaacaacca aactaccac caacgccacc gtcgcgacct 27780
ttcctctggg tctaatacca ctaccggagg tgagctccga ggtcgaccaa cctctgggat 27840
ttactacggc ccctgggagg tggtaggggt aatagcgcta ggcctagttg cgggtgggct 27900
tttggtcttc tgctacctat acctcccttg ctgttcgtac ttagtggtgc tgtgttgctg 27960
gtttaagaaa tggggaagat caccctagtg agctgcggtg tgctgggtggc ggtgggtgctt 28020
tcgattgtgg gactgggagg cgcggtgta gtgaaggaga aggccgatcc ctgcttgcat 28080
ttcaatcccg acaaatgccg gctgagtttt cagcccgatg gcaatcgggtg cgcggtgctg 28140
atcaagtgcg gatgggaatg cgagaacgtg agaatcgagt acaataacaa gactcggaac 28200
aatactctcg cgtccgtgtg gcagccggg gaccccgagt ggtacaccgt ctctgtcccc 28260
gggtgctgacg gctccccgcg caccgtgaat aatactttca tttttgcga catgtgcgac 28320
acgggtcatgt ggatgagcaa gcagtagcat atgtggcccc ccacgaagga gaacatcgtg 28380
gtcttctcca tcgcttacag cgtgtgcacg gcgctaata caagctatcgt gtgcctgagc 28440
attcacatgc tcatcgctat tcgccccaga aataatgccg aaaaagaaaa acagccataa 28500
cacgtttttt cacacacctt tttcagacca tggcctctgt taaatttttg cttttatttg 28560
ccagtctcat tgccgtcatt catggaatga gtaatgagaa aattactatt tacactggca 28620
ctaatacacac attgaaaggt ccagaaaaag ccacagaagt ttcattggtat tgttatttta 28680
atgaatcaga tgtatctact gaactctgtg gaaacaataa caaaaaaat gagagcatta 28740
ctctcatcaa gtttcaatgt ggatctgact taaccctaata taacatcact agagactatg 28800
taggtatgta ttatggaact acagcaggca tttcgacat ggaattttat caagtttctg 28860
tgtctgaacc caccacgcct agaatgacca caaccacaaa aactacacct gttaccacta 28920
tacagctcac taccaatggc tttcttgcca tgcttcaagt ggctgaaaat agcaccagca 28980
ttcaaccac cccaccagt gaggaatttc ccagatccat gattggcatt attgttgctg 29040

tagtggtgtg	catgttgatc	atcgccctgt	gcatggtgta	ctatgccttc	tgctacagaa	29100
agcacagact	gaacgacaag	ctggaacact	tactaagtgt	tgaattttaa	tttttttagaa	29160
ccatgaagat	cctaggcctt	ttagtttttt	ctatcattac	ctctgctcta	tgcaattctg	29220
acaatgagga	cgttactgtc	gttgctcgga	caaattatac	actaaaaggt	ccagcaaaag	29280
gtatgctttc	gtggtattgt	tggttcggaa	ctgacgagca	acagacagaa	ctttgcaatg	29340
ctcaaaaagg	caaaacctca	aattctaaaa	tctctaatta	tcaatgcaat	ggcactgact	29400
tagtattgct	caatgtcacg	aaagcatatg	ctggcagtta	cacctgccct	ggagatgatg	29460
ccgacaatat	gattttttac	aaagtggaag	tggttgatcc	cactactcca	ccgcccacca	29520
ccacaactac	tcataccaca	cacacagaac	aaacaccaga	ggcagcagaa	gcagagttgg	29580
ccttccaggt	tcacggagat	tcctttgctg	tcaatacccc	tacacccgat	cagcgggtgtc	29640
cggggctgct	cgtcagcggc	attgtcgggtg	tgctttcggg	attagcagtc	ataatcatct	29700
gcatgttcat	ttttgcttgc	tgctatagaa	ggctttaccg	acaaaaatca	gaccactgtc	29760
tgaacctcta	tgtttaattt	tttccagagc	catgaaggca	gttagcgctc	tagttttttg	29820
ttctttgatt	ggcattgttt	ttagtgctgg	gtttttgaaa	aatcttacca	tttatgaagg	29880
tgagaatgcc	actctagtgg	gcatcagtgg	tcaaaatgtc	agctggctaa	aataccatct	29940
agatgggtgg	aaagacattt	gcgattggaa	tgtcactgtg	tatacatgta	atggagttaa	30000
cctcaccatt	actaatgcca	ccaagatca	gaatggtagg	tttaagggcc	agagtttcac	30060
tagaaataat	gggtatgaat	cccataacat	gtttatctat	gacgtcactg	tcatacagaa	30120
tgagactgcc	accaccacac	agatgccac	tacacacagt	tctaccacta	ctaccatgca	30180
aaccacacag	acaaccacta	catcaactca	gcatatgacc	accactacag	cagcaaagcc	30240
aagtagtgca	gcgcctcagc	cccaggcttt	ggctttgaaa	gctgcacaac	ctagtacaac	30300
tactaggacc	aatgagcaga	ctactgaatt	tttgtccact	gtcgagagcc	acaccacagc	30360
tacctccagt	gccttctcta	gcaccgccaa	tctctcctcg	ctttcctcta	caccaatcag	30420
tcccgtact	actcccaccc	cagctcttct	ccccactccc	ctgaagcaaa	ctgaggacag	30480
cggcatgcaa	tggcagatca	ccctgctcat	tgtgatcggg	ttggtcatcc	tggccgtggt	30540
gctctactac	atcttctgcc	gccgcattcc	caacgcgcac	cgcaaaccgg	cctacaagcc	30600
catcgttatc	gggcagccgg	agccgcttca	gggtgaaggg	ggtctaagga	atcttctctt	30660
ctcttttaca	gtatggtgat	tgaactatga	ttcctagaca	attcttgatc	actattctta	30720
tctgcctcct	ccaagtctgt	gccaccctcg	ctctggtggc	caacgccagt	ccagactgta	30780
ttggggccctt	cgcctcctac	gtgctctttg	ccttcatcac	ctgcatctgc	tgctgtagca	30840
tagtctgcct	gcttatcacc	ttcttccagt	tcattgactg	gatctttgtg	cgcacgcct	30900
acctgcgcca	ccacccccag	taccgcgacc	agcgagtggc	gcggctgctc	aggctcctct	30960
gataagcatg	cgggctctgc	tacttctcgc	gcttctgctg	ttagtgctcc	ccgccccgt	31020
cgacccccgg	tccccactc	agtcccccga	agaggtccgc	aaatgcaa	tccaagaacc	31080
ctggaaattc	ctcaaatgct	accgccaaaa	atcagacatg	cttcccagct	ggatcatgat	31140

cattgggatc gtgaacattc tggcctgcac cctcatctcc tttgtgattt acccctgctt 31200
tgactttggt tggaactcgc cagaggcgct ctatctcccg cctgaacctg acacaccacc 31260
acagcaacct caggcacacg cactaccacc accacagcct aggccacaat acatgccccat 31320
attagactat gaggccgagc cacagcgacc catgctcccc gctattagtt acttcaatct 31380
aaccggcgga gatgactgac ccactggcca acaacaacgt caacgacctt ctcttgga 31440
tggacggccg cgcctcggag cagcgactcg cccaacttcg cattcgccag cagcaggaga 31500
gagccgtcaa ggagctgcag gacggcatag ccatccacca gtgcaagaaa ggcatcttct 31560
gcctggtgaa acaggccaag atctcctacg aggtcacccc gaccgaccat cgcctctcct 31620
acgagctcct gcagcagcgc cagaagttca cctgcctggt cggagtcaac cccatcgtca 31680
tcaccagca gtcgggcat accaaggggt gcatccactg ctctgcgac tccccgact 31740
gcgtccacac tctgatcaag accctctgcg gcctccgcga cctcctcccc atgaactaat 31800
cacccttta tccagtgaat taaatatcat attgatgatg atttaaataa aaaataatca 31860
tttgatttga aataaagata caatcatatt gatgatttga gttttaaaaa ataaagaatc 31920
acttacttga aatctgatac caggtctctg tccatgtttt ctgccaacac cacctcactc 31980
ccctcttccc agctctggta ctgcagaccc cggcgggctg caaacttcct ccacacgctg 32040
aaggggatgt caaattcctc ctgtccctca atcttcattt tatcttctat cagatgtcca 32100
aaaagcgcgt ccgggtggat gatgacttcg acccgtcta cccctacgat gcagacaacg 32160
caccgaccgt gcccttcac aacccccct tcgtctcttc agatggattc caagagaagc 32220
ccctgggggt gctgtccctg cgactggctg acccgtcac caccaagaac ggggaaatca 32280
ccctcaagct gggagagggg gtggacctcg actcctcggg aaaactcatc tccaacacgg 32340
ccaccaaggc cgccgcccct ctcagttttt ccaacaacac catttcctt aacatggata 32400
ccctcttta taccaaagat ggaaaattat ccttacaagt ttctccaccg ttaaacadat 32460
taaaatcaac cattctgaac acattagctg tagcttatgg atcaggttta ggactgagtg 32520
gtggcactgc tcttgagta cagttggcct ctccactcac ttttgatgaa aaaggaaata 32580
ttaaattaa cctagccagt ggtccattaa cagttgatgc aagtcgactt agtatcaact 32640
gcaaaagagg ggtcactgtc actacctcag gagatgcaat tgaaagcaac ataagctggc 32700
ctaaaggat aagatttgaa ggtaatggca tagctgcaaa cattggcaga ggattggaat 32760
ttggaaccac tagtacagag actgatgtca cagatgcata cccaattcaa gttaaattgg 32820
gtactggcct tacctttgac agtacaggcg ccattgttg tggaaacaaa gaggatgata 32880
aacttacatt atggaccaca gccgaccct cgccaaattg caaaatatac tctgaaaaag 32940
atgccaaact cacactttgc ttgacaaagt gtggaagtca aattctgggt actgtgactg 33000
tattggcagt gaataatgga agtctcaacc caatcacaaa cacagtaagc actgcactcg 33060
tctccctcaa gtttgatgca agtggagttt tgctaagcag ctccacatta gacaaagaat 33120
attggaactt cagaaagga gatgttacac ctgctgagcc ctataactaat gctatagggt 33180

ttatgcctaa cataaaggcc tatictaaaa acacatctgc agcttcaaaa agccatattg 33240
 tcagtcaagt ttatctcaat ggggatgagg ccaaaccact gatgctgatt attactttta 33300
 atgaaactga ggatgcaact tgcacctaca gtatcacttt tcaatggaaa tgggatagta 33360
 ctaagtacac aggtgaaaca cttgctacca gctccttcac cttctcctac atcgcccaag 33420
 aatgaacact gtatcccacc ctgcatgccca acccttccca cccactctg tctatggaaa 33480
 aaactctgaa gcacaaaata aaataaagtt caagtgtttt attgattcaa cagttttaca 33540
 ggattcgagc agttatTTTT cctccaccct cccaggacat ggaatacacc accctctccc 33600
 cccgcacagc cttgaacatc tgaatgccat tggatgatgga catgcttttg gtctccacgt 33660
 tccacacagt ttcagagcga gccagtctcg ggtcggctcag ggagatgaaa ccctccgggc 33720
 actcccgcat ctgcacctca cagctcaaca gctgaggatt gtcctcgggtg gtcgggatca 33780
 cggttatctg gaagaagcag aagagcggcg gtgggaatca tagtccgcga acgggatcgg 33840
 ccggtggtgt cgcacagggc cccgcagcag tcgctgccgc cgccgctccg tcaagctgct 33900
 gctcaggggg tccgggtcca gggactccct cagcatgatg cccacggccc tcagcatcag 33960
 tcgtctggtg cggcgggccc agcagcgcgt gcggatctcg ctgaggtcgc tgcagtacgt 34020
 gcaacacagg accaccaggt tgttcaacag tccatagttc aacacgctcc agccgaaact 34080
 catcgcggaagg aggatgctac ccacgtggcc gtcgtaccag atcctcaggt aaatcaagtg 34140
 gcgctccctc cagaacacgc tgcccacgta catgatctcc ttgggcatgt ggcgggtcac 34200
 cacctcccgg taccacatca ccctctggtt gaacatgcag ccccggtatga tcctgcggaa 34260
 ccacagggcc agcaccgccc cgcccgccat gcagcgaaga gaccccggtt cccggcaatg 34320
 gcaatggagg acccaccgct cgtaccctg gatcatctgg gagctgaaca agtctatggt 34380
 ggcacagcac aggcataatgc tcatgcatct cttcagcact ctgagctcct cgggggtcaa 34440
 aaccatatcc cagggcacgg ggaactcttg caggacagcg aaccccgag aacagggcaa 34500
 tcctcgcaca taacttacat tgtgcatgga cagggtatcg caatcaggca gcaccgggtg 34560
 atcctccacc agagaagcgc gggctctcgt ctctcacag cgtggttaagg gggccggccg 34620
 atacgggtga tggcgggacg cggctgatcg tgttcgcgac cgtgtcatga tgcagttgct 34680
 ttcggacatt ttcgtacttg ctgtagcaga acctggtccg ggcgctgcac accgatcgcc 34740
 ggcggcggtc ccggcgcttg gaacgctcgg tgttgaaatt gtaaaacagc cactctctca 34800
 gaccgtgcag cagatctagg gcctcaggag tgatgaagat cccatcatgc ctgatagctc 34860
 tgatcacatc gaccaccgtg gaatgggcca gaccagcca gatgatgcaa ttttggtggg 34920
 tttcggtgac ggcgggggag ggaagaacag gaagaacctat gattaacttt taatccaaac 34980
 ggtctcggag cacttcaaaa tgaaggctcg ggagatggca cctctcgccc ccgctgtggt 35040
 ggtggaaaat aacagccagg tcaaagggtga tacggttctc gagatgttcc acggtggctt 35100
 ccagcaaagc ctccacgcgc acatccagaa acaagacaat agcgaaagcg ggagggttct 35160
 ctaattcctc aatcatcatg ttacactcct gcaccatccc cagataattt tcatttttcc 35220
 agccttgaat gattcgaact agttcctgag gtaaatccaa gccagccatg ataaagagct 35280

cgcgagagc gccctccacc ggcattctta agcacaccct cataattcca agatattctg 35340
 ctcttggttc acctgcagca gattgacaag cggaatatca aaatctctgc cgcgatccct 35400
 aagctcctcc ctgagcaata actgtaagta ctctttcata tcctctccga aatttttagc 35460
 cataggacca ccaggaataa gattagggca agccacagta cagataaacc gaagtcctcc 35520
 ccagtgcgca ttgccaaatg caagactgct ataagcatgc tggctagacc cggtgatatc 35580
 ttccagataa ctggacagaa aatcacccag gcaattttta agaaaatcaa caaaagaaaa 35640
 atcctccagg tgcacgttta gagcctcggg aacaacgatg aagtaaattgc aagcgggtgcg 35700
 ttccagcatg gttagtttagc tgatctgtaa aaaacaaaaa ataaaacatt aaaccatgct 35760
 agcctggcga acaggtgggt aaatcgttct ctccagcacc aggcaggcca cgggggtctcc 35820
 ggcgcgaccc tcgtaaaaat tgtcgtctatg attgaaaacc atcacagaga gacgttccccg 35880
 gtggccggcg tgaatgattc gacaagatga atacaccccc ggaacattgg cgtccgcgag 35940
 tgaaaaaaag cgcccagga agcaataagg cactacaatg ctgagtctca agtccagcaa 36000
 agcgatgcca tgcggatgaa gcacaaaatc ctgaggtgcg taaaaaatgt aattactccc 36060
 ctctgcaca ggcagcgaag ccccgatcc ctccagatac acatacaaag cctcagcgtc 36120
 catagcttac cgagcagcag cacacaacag gcgcaagagt cagagaaagg ctgagctcta 36180
 acctgtccac ccgtctctctg ctcaatatat agcccagatc tacttgacg taaaggccaa 36240
 agtctaaaaa taccgcgcaa ataatacacac acgcccagca cacgcccaga aaccgggtgac 36300
 aactcaaaa aaatacgcg acttctctcaa acgcccacaa tgccgtcatt tccgggttcc 36360
 cacgctacgt catcggaatt cgactttcaa attccgtcga ccgttaaaaa cgtcacccgc 36420
 cccgccccta acggtcgccc gtctctcggc caatcacctt cctccctccc caaattcaaa 36480
 cagctcattt gcatattaac gcgcaccaa agtttgaggt atattattga tgatg 36535

<210> 4
 <211> 34264
 <212> DNA
 <213> simian adenovirus SV-1

<400> 4
 tccttattct ggaaacgtgc caatatgata atgagcgggg aggagcgagg cggggccggg 60
 gtgacgtgcg gtgacgtggg gtgacgcggg gtggcgcgag ggcggggcgg gagggtgggag 120
 gcgcttagtt ttacgtatg cggaaggagg ttttataacc gaagttgggt aatttgggag 180
 tatacttgta agttttgtgt aatttggcgc gaaaaccggg taatgaggaa gttgaggtta 240
 atatgtactt tttatgactg ggcggaattt ctgctgatca gcagtgaact ttgggcgctg 300
 acggggagggt ttcgctacgt ggcagtagca cgagaaggct caaagggtccc atttattgta 360
 ctctcagcg ttttcgctgg gtatttaaac gctgtcagat catcaagagg cactccttga 420
 gtgccggcga gtagagtttt ctctccgcg ctgccgcgat gaggctgggt cccgagatgt 480
 acggtgtttt ctgcagcgag acggcccgga actcagatga gctgcttaat acagatctgc 540
 tggatgttcc caactcgctt gtggcttcgc ctccgtcgct tcatgatctt ttcgatgtgg 600

aagtggatcc accgcaagat cccaacgagg acgcggtaaa cagtatgttc cctgaatgtc	660
tgtttgaggc ggctgaggag ggttctcaca gcagtgaaga gagcagacgg ggagaggaac	720
tggacttgaa atgctacgag gaatgtctgc cttctagcga ttctgaaacg gaacagacag	780
ggggagacgg ctgtgagtcg gcaatgaaaa atgaacttgt attagactgt ccagaacatc	840
ctggtcatgg ctgccgtgcc tgtgcttttc atagaaatgc cagcggaaat cctgagactc	900
tatgtgctct gtgttatctg cgccttacca gcgattttgt atacagtaag taaagtgttt	960
tcattggcgt acggtagggg attcgttgaa gtgctttgtg acttattatg tgtcattatt	1020
tctaggtgac gtgtccgacg tggaagggga aggagataga tcaggggctg ctaattctcc	1080
ttgcactttg ggggctgtgg ttccagttgg catttttaaa ccgagtgggtg gaggagaacg	1140
agccggagga gaccgagaat ctgagagccg gcctggaccc tccagtggaa gactaggtgc	1200
tgaggatgat cctgaagagg ggactagtgg ggggtgctagg aaaaagcaaa aaactgagcc	1260
tgaacctaga aactttttga atgagttgac tgtaagccta atgaatcggc agcgtcctga	1320
gacggtgttt tggactgagt tggaggatga gttcaagaag ggggaattaa acctcttgta	1380
caagtatggg tttgagcagt tgaaaactca ctggttggag ccgtgggagg atatggaaat	1440
ggctctagac acctttgcta aagtggctct gcggccggat aaagtgttaca ctattcgccg	1500
cactgttaat ataaaaaaga gtgtttatgt tatcgcccat ggagctctgg tgcaggtgca	1560
gacccagac cgggtggctt tcaattgcgg catgcagagt ttgggccccg ggggtgatagg	1620
tttgaatgga gttacatttc aaaatgtcag gtttactggt gatgatttta atggctctgt	1680
gtttgtgact agcaccagc taacctcca cggtgtttac ttttttaact ttaacaatac	1740
atgtgtggag tcatggggtg ggggtgtctct gaggggctgc agttttcatg gttgctggaa	1800
ggcgggtggg ggaagaatta aaagtgtcat gtctgtgaag aaatgcatat ttgaacgctg	1860
tgtgatagct ctagcagtag aggggtacgg acggatcagg aataacgccg catctgagaa	1920
tggatgtttt cttttgctga aaggtagcgc cagcgtaag cataatatga tttgcggcag	1980
cggcctgtgc ccctcgcagc tcttaacttg cgcagatgga aactgtcaca ccttgcgcac	2040
cgtgcacata gtgtcccact cgcgccgcac ctggccaaca tttgagcaca atatgctcat	2100
gcgttgccg gttcacctag gtgctagacg cggcgtgttt atgccttatc aatgtaactt	2160
tagtcatact aagattttgc tggaactga ttccttcct cgagtatgtt tcaatggggt	2220
gtttgacatg tcaatggaac tttttaagt gataagatat gatgaaacca agtctcgttg	2280
tcgctcatgt gaatgcggag ctaatcattt gaggttgat cctgtaacct tgaacgttac	2340
cgaggagctg aggacggacc accacatgct gtcttgctg cgtaccgact atgaatccag	2400
cgatgaggag tgaggtgagg ggcggagcca caaagggtat aaaggggcat gagggggtgg	2460
cgcgggtgtt caaaatgagc gggacgacgg acggcaatgc gtttgagggg ggagtgttca	2520
gcccatatct gacatctcgt cttccttcct gggcaggagt tcgtcagaat gtagtgggct	2580
ccaccgtgga cggacggccg gtcgcccctg caaattccgc caccctcacc tatgccaccg	2640

tgggatcatc	gttggacact	gccgcggcag	ctgccgcttc	tgctgccgct	tctactgctc	2700
gcggcatggc	ggctgatttt	ggactatata	accaactggc	cactgcagct	gtggcgtctc	2760
ggtctctggt	tcaagaagat	gccctgaatg	tgatcttgac	tcgcctggag	atcatgtcac	2820
gtcgcctgga	cgaactggct	gcgcagatat	cccaagctaa	ccccgatacc	gcttcagaat	2880
cttaaaataa	agacaaacaa	at ttgttgaa	aagtaaaatg	gctttatttg	tttttttttg	2940
ctcggtaggc	tcgggtccac	ctgtctcggt	cg ttaaggac	tttgtgtatg	ttttccaaaa	3000
cacggtacag	atgggcttgg	atgttcaagt	acatgggcat	gaggccatct	ttgggggtgga	3060
gataggacca	ctgaagagcg	tcatgttccg	gggtgggtatt	gtaaatcacc	cagtcgtagc	3120
aggggtttttg	agcgtggaac	tggaaatatgt	ccttcaggag	caggcta atg	gccaagggtga	3180
gacccttagt	gtaggtgttt	acaaagcggt	tgagctggga	gggatgcatg	cggggggaga	3240
tgatatgcat	cttggcttgg	at ttgtgaggt	tagctatgtt	accacccagg	tctctgcggg	3300
ggttcatgtt	atgaaggacc	accagcacgg	tatagccagt	gcattttgggg	aacttgtcat	3360
gcagtttggga	ggggaaggcg	tggagaattt	tagatacccc	cttgtgcccc	cctagggtttt	3420
ccatgcactc	atccataata	atggcaatgg	gaccctggc	ggccgcttta	gcaaacacgt	3480
tttggggggtt	ggaaacatca	tagttttgct	ctagagttag	ctcatcatag	gccatcttta	3540
caaagcgggg	taggaggggtg	cccgactggg	ggatgatagt	tccatctggg	cctggagcgt	3600
agttgccttc	acagatctgc	atctcccagg	ccttaatttc	cgagggggggg	atcatgtcca	3660
cctgggggggc	gataaaaaac	acggtttctg	gcgggggggtt	aatgagctgg	gtggaaagca	3720
agttacgcaa	cagctgggat	ttgccgcaac	cggtgggacc	gtagatgacc	ccgatgacgg	3780
gttgacagctg	gtagttcaga	gaggaacagc	tgccgtcggg	gcgcaggagg	ggagctacct	3840
cattcatcat	gcttctgaca	tg tttatttt	cactcactaa	gtttttgcaag	agcctctccc	3900
caccagggga	taagagttct	tccaggctgt	tgaagtgttt	cagcggtttc	aggccgtcgg	3960
ccatgggcat	cttttcaagc	gactgacgaa	gcaagtacag	tcgggtcccag	agctcgggtga	4020
cgtgctctat	ggaatctcga	tccagcagac	ttcttggttt	cgggggttgg	gccgactttc	4080
gctgtagggc	accagccgggt	gggcgtccag	ggccgcgagg	gttctgtcct	tccagggtct	4140
cagcgttcgg	gtgaggggtg	tctcgggtgac	ggtgaaggga	tgagccccgg	gctgggcgct	4200
tgcgaggggtg	cgcttcaggc	tcatcctgct	ggtgctgaag	cgggcgtcgt	ctccctgtga	4260
gtcggccaga	tagcaacgaa	gcatgaggtc	gtagctgagg	gactcggccg	cgtgtccctt	4320
ggcgcgcagc	tttcccttgg	aaacgtgctg	acatttggtg	cagtgcagac	acttgagggc	4380
gtagagtttt	ggggccagga	agaccgactc	gggcgagtag	gcgtcggctc	cgactgagc	4440
gcagacggtc	tcgcactcca	ccagccacgt	gagctcgggt	ttagcgggat	caaaaaccaa	4500
gttgccctcca	ttttttttga	tgcgtttctt	accttgctgc	tccatgagtc	tgtgtcccg	4560
ttccgtgaca	aaaaggctgt	cggtatcccc	gtagaccgac	ttgagggggc	gatcttccaa	4620
aggtgttccg	aggtcttccg	cgtaaggaa	ctgggaccac	tccgagacaa	aggctcgggt	4680
ccaggctaac	acgaaggagg	cgatctgcga	gggttatctg	tcgttttcaa	tgaggggggtc	4740

caccttttcc	agggtgtgca	gacacaggtc	gtcctcctcc	gcgtccacga	agggtgattgg	4800
cttgtaagtg	taggtcacgt	gacccgcacc	cccccaagg	gtataaaagg	gggcgtgccc	4860
actctccccg	tcacttttctt	ccgcatcgct	gtggaccaga	gccagctggt	cgggtgagta	4920
ggccctctca	aaagccggca	tgatttcggc	gctcaagttg	tcagtttcta	caaacgaggt	4980
ggatttgata	ttcacgtgcc	ccgcggcgat	gcttttgatg	gtggaggggt	ccatctgac	5040
agaaaacacg	atctttttat	tgtcaagttt	ggtggcgaaa	gacccgtaga	gggcgttgga	5100
aagcaacttg	gcgatggagc	gcagggctctg	atttttctcc	cgatcgcccc	tctccttggc	5160
ggcgatggtg	agttgcacgt	actcgcgggc	cacgcaccgc	cactcgggga	acacggcggt	5220
gcgctcgctg	ggcaggatgc	gcacgcgcca	gccgcggttg	tgacgggtga	tgagggtccac	5280
gctggtggcc	acctccccgc	ggaggggctc	gttggtccaa	cacaatcgcc	ccccctttct	5340
ggagcagaac	ggaggcagg	gatctagcaa	gttggcgggc	gggggggtcgg	cgtcgatggt	5400
aaatatgccg	ggtagcagaa	ttttattaaa	ataatcgatt	tcggtgtccg	tgtcttgcaa	5460
cgcgtcttcc	cacttcttca	ccgccagggc	cctttcgtag	ggattcagg	gcggtcccca	5520
gggcatgggg	tgggtcagg	ccgaggcgta	catgccgcag	atgtcgtaca	cgtacagggg	5580
ctccctcaac	accccgatgt	aagtggggta	acagcgcccc	ccgcggatgc	tggctcgcac	5640
gtagtcgtac	atctcgtgag	agggagccat	gagcccgtct	cccaagtggg	tcttggtggg	5700
tttttcggcc	cggtagagga	tctgcctgaa	gatggcgtgg	gagttggaag	agatagtggg	5760
gcgttggaag	acgttaaagt	tggctccggg	cagtcccacg	gagtcttgga	tgaactgggc	5820
gtaggattcc	cggagcttgt	ccaccagggc	tgcggttacc	agcacgtcga	gagcgagta	5880
gtccaacgtc	tcgcggacca	ggttgtaggc	cgtctcttgt	tttttctccc	acagttcgcg	5940
attgaggagg	tattcctcgc	ggtctttcca	gtactcttcg	gcgggaaatc	ctttttcgtc	6000
cgctcggtaa	gaacctaac	tgtaaaattc	gttcacggct	ttgtatggac	aacagccttt	6060
ttctaccggc	agggcgtacg	cttgagcggc	ctttctgaga	gaggtgtggg	tgagggcgaa	6120
ggtgtcccg	accatcactt	tcagggtactg	atgtttgaag	tccgtgtcgt	cgcaggcgcc	6180
ctgttccac	agcgtgaagt	cgggtgcgctt	tttctgcctg	ggattgggga	gggcgaatgt	6240
gacgtcggtta	aagaggattt	tcccggcgcg	gggcatgaag	ttgcgagaga	tcctgaagg	6300
tccgggcacg	tccgagcggt	tgttgatgac	ttgcgcggcc	aggacgatct	cgtcgaagcc	6360
gttgatggtg	tggcccacga	tgtaaagtgc	gataaagcgc	ggctgtccct	tgagggccgg	6420
cgcttttttc	aactcctcgt	agggtgagaca	gtccggcgag	gagagaccca	gctccgcccc	6480
ggcccagtcg	gagagctgag	ggttagccgc	gaggaaagag	ctccacaggt	caagggctag	6540
cagagtttgc	aagcggctgc	ggaactcgcg	aaactttttc	cccacggcca	ttttctccgg	6600
cgtcaccacg	tagaaagtgc	aggggcggtc	gttccagacg	tcccatcgga	gctctagggc	6660
cagctcgag	gcttgacgaa	cgaggggtctc	ctcgcccgag	acgtgcatga	ccagcatgaa	6720
gggtaccaac	tgtttccga	acgagcccat	ccatgtgtag	gtttctacgt	cgtaggtgac	6780

aaagagccgc	tgggtgcgcg	cgtgggagcc	gatcgggag	aagctgatct	cctgccacca	6840
gttgaggaa	tgggtgttga	tgtggtgaaa	gtagaagtcc	cgccggcgca	cagagcattc	6900
gtgctgatgt	ttgtaaaagc	gaccgcagta	gtcgcagcgc	tgcacgctct	gtatctcctg	6960
aatgagatgc	gcttttcgcc	cgcgaccag	aaaccggagg	gggaagtga	gacgggggct	7020
tgggtggggcg	gcatcccctt	cgcttggcg	gtgggagtct	gcgtctgcgc	cctccttctc	7080
tgggtggacg	acggtgggga	cgacgacgcc	ccgggtgccg	caagtccaga	tctccgccac	7140
ggaggggcg	aggcgttgca	ggaggggacg	cagctgcccc	ctgtccaggg	agtcgagggc	7200
ggccgcgctg	aggtcggcg	gaagcgtttg	caagttcact	ttcagaagac	cggtaaagac	7260
gtgagccagg	tgcacatggt	acttgatttc	caggggggtg	ttggaagagg	cgtccacggc	7320
gtagaggagg	ccgtgtccgc	gcggggccac	caccgtgccc	cgaggagggt	ttatctcact	7380
cgtcgagggc	gagcgccggg	gggtagaggc	ggctctgcgc	cggggggcag	cggaggcagt	7440
ggcacgtttt	cgtgaggatt	cggcagcggt	tgatgacgag	cccggagact	gctggcggtg	7500
gcgacgacgc	ggcggttgag	gtcctggatg	tgccgtctct	gcgtgaagac	caccggcccc	7560
cgggtcctga	acctgaaaga	gagttccaca	gaatcaatgt	ctgcatcggt	aacggcgggc	7620
tgcctgagga	tctcctgtac	gtcgcccag	ttgtcttgat	aggcgatctc	ggccatgaac	7680
tgctccactt	cttcctcgcg	gaggtcgccg	tggcccgtct	gctccacggt	ggcggccagg	7740
tcgttgagga	tgcgacgcat	gagttgagag	aaggcggtga	ggccgttctc	gttccacacg	7800
cggctgtaca	ccacgtttcc	gaaggagtcg	cgcgctcgca	tgaccacctg	ggccacgttg	7860
agttccacgt	ggcgggagaa	gacggcgtag	tttctgaggc	gctggaagag	gtagttgagc	7920
gtggtggcga	tgtgctcgca	gacgaagaag	tacatgatcc	agcgccgag	ggcatctctg	7980
ttgatgtctc	cgatggcttc	gagacgctcc	atggcctcgt	agaagtcgac	ggcgaagttg	8040
aaaaattggg	agttgcgggc	ggccaccgtg	agttcttctt	gcaggaggcg	gatgagatcg	8100
gcgaccgtgt	cgcgacctc	ctgctcgaaa	gcgccccgag	gcgcctctgc	ttcttcctcc	8160
ggctcctcct	cttcaggggg	cacgggttcc	tccggcagct	ctgcgacggg	gacggggcg	8220
cgacgtcgtc	gtctgaccgg	caggcggtcc	acgaagcgct	cgatcatttc	gccgcgccgg	8280
cgacgcatgg	tctcggtgac	ggcggtccg	tttctcgag	gtcgcagttc	gaagacgccg	8340
ccgcgcagag	cgccccgtg	cagggagggt	aagtggttag	ggccgtcggg	cagggacacg	8400
gcgctgacga	tgcattttat	caattgctgc	gtaggcactc	cgtgcaggga	tctgagaacg	8460
tcgaggtcga	cgggatccga	gaacttctct	aggaaagcgt	ctatccaatc	gcagtcgcaa	8520
ggtaagctga	ggacgggtggg	ccgctggggg	gcgtccgcgg	gcagttggga	ggtgatgctg	8580
ctgatgatgt	aattaaagta	ggcggctctc	aggcggcgga	tgggtggcgag	gaggaccacg	8640
tctttggggc	cggcctgttg	aatgcgcagg	cgctcgggca	tgccccaggc	ctcgctctga	8700
cagcgacgca	ggtctttgta	gtagtcttgc	atcagtctct	ccaccggaac	ctctgcttct	8760
cccctgtctg	ccatgcgagt	cgagccgaac	ccccgcaggg	gctgcagcaa	cgctaggtcg	8820
gccacgaccc	tctcgccag	cacggcctgt	tggatctgcg	tgagggtggt	ctggaagtcg	8880

tccaggtcca	cgaagcggtg	ataggccccc	gtgttgatgg	tgtaggtgca	gttggccatg	8940
acggaccagt	tgacgacttg	catgccgggt	tgggtgatct	ccgtgtactt	gaggcgcgag	9000
taggcgcggg	actcgaacac	gtagtcgttg	catgtgcgta	ccagatactg	gtagccaacc	9060
aggaagtggg	gaggcggttc	tcggtacagg	ggccagccga	ctgtggcggg	ggcgccgggg	9120
gacaggtcgt	ccagcatgag	gcatggtag	tggtagatgt	agcgggagag	ccaggtgatg	9180
ccggccgagg	tggtcgcggc	cctggtgaat	tcgcggacgc	ggttccagat	gttgcgcagg	9240
gggcgaaagc	gtcccatggt	gggcacgctc	tgccccgtga	ggcgggcgca	atcttgtagc	9300
ctctagatgg	aaaaaagaca	gggcggtcat	cgactccctt	ccgtagctcg	gggggtaaag	9360
tcgcaagggg	gcggcgggcg	ggaaccccgg	ttcgagaccg	gccggatccg	ccgctcccga	9420
tgcgcctggc	cccgcattca	cgacgtccgc	gtcgagaccc	agccgcgacg	ctccgcccc	9480
atacggaggg	gagtcttttg	gtgttttttc	gtagatgcat	ccggtgctgc	ggcagatgcg	9540
acctcagacg	cccaccacca	ccgccgcggc	ggcagtaaac	ctgagcggag	gcggtgacag	9600
ggaggaggag	gagctggctt	tagacctgga	agagggagag	gggctggccc	ggctgggagc	9660
gccgtcccca	gagagacacc	ctagggttca	gctcgtgagg	gacgccaggc	aggcttttgt	9720
gccgaagcag	aacctgttta	gggaccgcag	cggtcaggag	gcggaggaga	tgcgcgattg	9780
caggtttcgg	gcgggtagag	agctgagggc	gggcttcgat	cgggagcggc	tcctgagggc	9840
ggaggatttc	gagcccagc	agcgttcttg	ggtgagccc	gcccgcgctc	acgtctcggc	9900
ggccaacctg	gtgagcgcgt	acgagcagac	ggtgaacgag	gagcgcaact	tccaaaagag	9960
ctttaacaat	cacgtgagga	ccctgatcgc	gagggaggag	gtgaccatcg	ggctgatgca	10020
tctgtgggac	ttcgtggagg	cctacgtgca	gaaccggcc	agcaaacctc	tgacggccca	10080
gctgttcctg	atcgtgcagc	acagccgcga	caacgagacg	ttccgcgacg	ccatgttgaa	10140
catcgcgagg	cccgagggtc	gctggctctt	ggatctgatt	aacatcctgc	agagcatcgt	10200
ggtgcaggag	aggggcctca	gcttagcgga	caaggtggcg	gccattaact	attcgatgca	10260
gagcctgggg	aagtcttacg	ctcgcaagat	ctacaagagc	ccttacgtgc	ccatagacaa	10320
ggagggtgaag	atagacagct	tttacatgcg	catggcgctg	aagggtgctga	cgctgagcga	10380
cgacctcggc	gtgtaccgta	acgacaagat	ccacaaggcg	gtgagcgcca	gccgccggcg	10440
ggagctgagc	gacaggggagc	tgatgcacag	cctgcagagg	gcgctggcgg	gcgccgggga	10500
cgaggagcgc	gaggcttact	tcgacatggg	agccgatctg	cagtggcgtc	ccagcgcgcg	10560
cgccttgagg	gcggcgggct	accccagcga	ggaggatcgg	gacgatttgg	aggaggcagg	10620
cgagtacgag	gacgaagcct	gaccgggcag	gtgttggttt	agatgcagcg	gccggcggac	10680
ggggccaccg	cggatcccgc	acttttgcca	tccatgcaga	gtcaaccttc	gggcgtgacc	10740
gcctccgatg	actgggcggc	ggccatggac	cgcattatgg	cgctgactac	ccgcaacccc	10800
gaggctttta	gacagcaacc	ccaggccaac	cgtttttcgg	ccatcttgga	agcggtggtg	10860
ccctcccgcga	ccaacccac	acacgagaaa	gtcctgacta	tcgtgaacgc	cctggtagac	10920

agcaaggcca tccgccgcga cgaggcgggc ttgatttaca acgctctgct ggaacggggtg 10980
gcgcgctaca acagcactaa cgttcagacc aatctggatc gcctcaccac cgacgtgaag 11040
gaggcgctgg ctcaagaagga gcggtttctg agggacagca atctgggctc tctgggtggca 11100
ctcaacgcct tcctgagcac gcagccggcc aacgtgcccc gcgggcagga ggactacgtg 11160
agcttcatca gcgctctgag gctgctgggtg tccgaggtgc cccagagcga ggtgtatcag 11220
tctgggcccgg attacttctt ccagacgtcc cgacagggtg tgcaaacggg gaacctgact 11280
caggccttta aaaacttgca aggcattgtg ggcgttaagg ccccggtggg cgatcgagcc 11340
accatctcca gtctgctgac ccccaacact cgcctgctgc tgctcttgat cgcgccgttc 11400
accaacagta gcactatcag ccgtgactcg tacctgggtc atctcatcac tttgtaccgc 11460
gaggccatcg gtcaggctca gatcgacgag cacacatatc aggagatcac taacgtgagc 11520
cgggccctgg gtcaggaaga taccggcagc ctggaagcca cgttgaactt tttgctaacc 11580
aaccggaggc aaaaaataacc ctccagttt acgttaagcg ccgaggagga gaggattctg 11640
cgatacgtgc agcagtcctg gagtctgtac ttgatgcggg agggcgccac cgcttccacg 11700
gctttagaca tgacggctcg gaacatggaa ccgtcctttt actccgcccc ccggccgttc 11760
attaaccgtc tgatggacta cttccatcgc gcggccgcca tgaacgggga gtacttcacc 11820
aatgccatcc tgaatccgca ttggatgccc ccgtccggct tctacaccgg cgagtttgac 11880
ctgcccgaag ccgacgacgg ctttctttgg gacgacgtgt ccgacagcat tttcacgccg 11940
ggcaatcgcc gattccagaa gaaggagggc ggagacgagc tccccctctc cagcgtggag 12000
gcggcctcta ggggagagag tccctttccc agtctgtctt ccgccagcag tggtcgggta 12060
acgcgcccgc gggtgcccgg ggagagcgac tacctgaacg accccttgct gcggccggct 12120
aggaagaaaa atttcccca caacgggggtg gaaagcttg tggataaaat gaatcgttgg 12180
aagacctacg cccaggagca gcgggagtg gaggacagtc agccgcgacc gctggttccg 12240
ccgcactggc gtcgtcagag agaagacctg gacgactccg cagacgatag tagcgtgttg 12300
gacctgggag ggagcggagc caacccttt gctcacttgc aaccaagg gcggtccagt 12360
cgcctctact aataaaaaag acgcggaaac ttaccagagc catggccaca gcgtgtgtcc 12420
tttcttctc tctttcttcc tcggcgcggc agaatagagaa gagcgggtgag agtcacgccg 12480
gcggcgatat aggggtccgc cccttcttac gaaagcgtga tgggatcagc gaacgtgccg 12540
gccacgctgg aggcgcctta cgttcctccc agatacctgg gacctacgga gggcagaaac 12600
agcatccgtt actccgagct ggcacccctg tacgatacca ccaagggtgta cctgggtggac 12660
aacaagtcgg cggacatcgc ctccctgaat tatcaaaacg atcacagcaa ttttctgact 12720
accgtgggtg agaacaatga cttcaccctc acggaggcgg gcacgcagac cattaacttt 12780
gacgagcgtt cccgctgggg cggtcagctg aaaccatcc tgcacaccaa catgcccac 12840
atcaacgagt tcatgtccac caacaagttc agggccaggc tgatgggttaa aaaggctgaa 12900
aaccagcctc ccgagtacga atggtttgag ttcaccattc ccgagggcaa ctattccgag 12960
accatgacta tcgatctgat gaacaatgcg atcgtggaca attacctgca agtggggagg 13020

cagaacgggg tattggaaa cgatatcggc gtaaaatttg ataccagaaa cttccgactg 13080
gggtgggatc ccgtgaccaa gctgggtgatg ccaggcgtgt acaccaacga ggcttttcac 13140
cccgacatcg tgctgctgcc ggggtgcggt gtggacttca ctcagagccg tttgagtaac 13200
ctgttaggga tcagaaagcg ccgccccttc caagagggtt ttcagatcat gtatgaggac 13260
ctggaaggag gtaacattcc aggtttgcta gacgtgccgg cgtatgaaga gagtgttaaa 13320
caggcggagg cgcagggacg agagattcga ggcgacacct ttgccacgga acctcacgaa 13380
ctggtataaa aacctctgga acaagacagt aaaaaacgga gttacaacat tatatccggc 13440
actatgaata ccttgtagcg gagctgggtt ctggcttaca actacgggga tcccgaag 13500
ggagtgaagt catggacat actcaccacc acggacgtga cctgcggctc gcagcaagt 13560
tactggtccc tgccggatat gatgcaagac ccggtcacct tccgcccctc caccgaagt 13620
agcaacttcc cgggtggtgg caccgagctg ctgcccgtcc atgccaagag cttctacaac 13680
gaacaggccg tctactcgca actcattcgc cagtcaccgc cgcttaccca cgtgttcaat 13740
cgctttcccg agaaccagat tctggtgcgc cctcccgctc ctaccattac caccgtcagt 13800
gaaaacgttc ccgccctcac agatcacgga accctgccgc tgcgcagcag tatcagtga 13860
gttcagcgcg tgaccatcac cgacgccaga cgtcgaacct gtccctacgt ttacaaagct 13920
cttggcgtag tggctcctaa agtgctctct agtcgcacct tctaaacatg tccatcctca 13980
tctctcccga taacaacacc ggctggggac tgggctccgg caagatgtac ggcggagcca 14040
aaaggcgctc cagtcagcac ccagttcgag ttcggggcca cttccgtgct ccctggggag 14100
cttacaagcg aggactctcg ggccgaacgg cggtagacga taccatagat gccgtgattg 14160
ccgacgcccg ccggtacaac ccgggaccgg tcgctagcgc cgctccacc gtggattccg 14220
tgatcgacag cgtggtagct ggcgctcggg cctatgctcg ccgcaagagg cggctgcac 14280
ggagacgtcg cccaccgcc gccatgctgg cagccagggc cgtgctgagg cgggcccgga 14340
gggtaggcag aagggtctat gcgcgcgtg ccgccaacgc cgccgccggg agggcccgcc 14400
gacaggctgc ccgccaggct gctgccgcca tcgctagcat ggccagacct aggagaggga 14460
acgtgtactg ggtgcgcgat tctgtgacgg gagtccgagt gccggtgcgc agccgacctc 14520
cccgaagtta gaagatcaa gctgcgaaga cggcgggtact gagtctccct gttgttatca 14580
gccaacatg agcaagcgca agtttaaga agaactgctg cagacgctgg tgcctgagat 14640
ctatggccct ccggacgtga agcctgacat taagccccgc gatataagc gtgttaaaaa 14700
gcgggaaaag aaagaggaa tcgcggtggt agacgatggc ggagtggaa ttattaggag 14760
tttcgccccg cgacgcaggg ttcaatggaa agggcggcgg gtacaacgcg ttttgaggcc 14820
gggcaccgcg gtagttttta cccggggaga gcggctcggc gttaggggtt tcaaaaggca 14880
gtacgacgag gtgtacggcg acgaggacat attggaacag gcggctcaac agatcggaga 14940
atttgcttac ggaaagcgtt cgcgtcgca agacctggcc atcgctttag acagcggcaa 15000
ccccacgccc agcctcaaac ctgtgacgct gcagcagggt ctccccgtga gcgccagcac 15060

ggacagcaag aggggaataa aaagagaaat ggaagatctg cagcccacca tccagctcat 15120
ggtccctaaa cggcagaggc tggaagaggt cctggagaaa atgaaagtgg acccaagcat 15180
agagccggac gtcaaagtca ggccgatcaa agaagtggcc cctgggtctcg ggggtgcagac 15240
ggtggatatc cagatccccg tcacgtcagc ttcgaccgcc gtggaagcca tggaaacgca 15300
aacggaaacc cctgccgcga tcggtaccag ggaagtggcg ttgcaaaccg acccctggta 15360
cgaatacgcc gcccctcggc gtcagaggcg acccgctcgt tacggccccg ccaacgccat 15420
catgccagaa tatgcgctgc atccgtctat cctgcccacc cccggctacc ggggagtgac 15480
gtatcgcccc tcaggaaccc gccgccgaac ccgtcgccgc cgccgctccc gtcgtgctct 15540
ggccccctg tcggtgcgcc gcgtaacacg ccggggaaaag acagttacca ttcccaacc 15600
gcgctaccac cctagcatcc tttaatgact ctgccgtttt gcagatggct ctgacttgcc 15660
gcgtgcgctt tcccgttccg cactatcgag gaagatctcg tcgtaggaga ggcatggcgg 15720
gtagtggctg ccggcgggct ttgcgaggc gcatgaaagg cggaaatttta cccgctctga 15780
tacccataat cgccgccgcc atcgggtgcc taccggcgt cgcttcagtg gccttgcaag 15840
cagctcgtaa taaataaacg aaggcttttg cacttatgtc ctggctcctga ctattttatg 15900
cagaaagagc atggaagaca tcaattttac gtcgctggct ccgcggcacg gctcgcgcc 15960
gctcatgggc acctggaacg acatcggcac cagtcagctc aacgggggcg ctttcaattg 16020
ggggagcctt tggagcggca ttaaaaactt tggtccacg attaaatcct acggcagcaa 16080
agcctggaac agtagtgctg gtcagatgct ccgagataaa ctgaaggaca ccaacttcca 16140
agaaaaagtg gtcaatgggg tggtgaccgg catccacggc gcggtagatc tcgccaacca 16200
agcgggtgcag aaagagattg acaggcgttt ggaaagctcg cgggtgccgc cgagagagg 16260
ggatgagggtg gaggtcgagg aagtagaagt agaggaaaag ctgccccgc tggagaaagt 16320
tcccgggtgcg cctccgagac cgcagaagcg acccaggcca gaactagaag aaactctggt 16380
gacggagagc aaggagcctc cctcgtaga gcaagccttg aaagagggcg cctctccacc 16440
ctaccaatg acaaaaccga tcgcgcttat ggctcggccg gtgtacggga aggactacaa 16500
gcctgtcacg ctagagctcc ccccgccgcc accgccccc cccacgcgcc cgaccgttcc 16560
cccccccctg ccggctccgt cggcgggacc cgtgtccgca cccgtcgccg tgctctgccc 16620
agccgcccgc ccagtggccg tggccactgc cagaaacccc agaggccaga gaggagccaa 16680
ctggcaaagc acgctgaaca gcatcgtggg cctgggagtg aaaagcctga aacgccgccg 16740
ttgctattat taaaagtgtg gctaaaaaat ttcccgttgt atacgcctcc tatgttaccg 16800
ccagagacgc gtgactgtcg ccgcgagcgc cgctttcaag atggccaccc catcgatgat 16860
gccgcagtgg tcttacatgc acatcgccgg gcaggacgcc tcggagtacc tgagccccgg 16920
tctcgtgcag ttcgcccgcg ccaccgacac ctacttcagc ttgggaaaca agtttagaaa 16980
ccccaccgtg gccccacccc acgatgtaac cacggaccgc tcgcaaaggc tgaccctgcy 17040
ttttgtgccc gtagaccggg aggacaccgc gtactcttac aaagtgcgct acacgctggc 17100
cgtaggggac aaccgagtgc tggacatggc cagcacctac ttgacatcc ggggagtgct 17160

ggatcgcggt cccagtttta agccctactc gggtagccgcg tacaattccc tggctcccaa 17220
 gggcgctccc aaccctgcag aatggacgaa ttcagacagc aaagttaaag tgagggcaca 17280
 ggcgcctttt gttagctcgt atggtgctac agcgattaca aaagagggta ttcaggtggg 17340
 agtaacctta acagactccg gatcaacacc acagtatgca gataaaacgt atcagcctga 17400
 gccgcaaatt ggagaactac agtggaacag cgatgttgga accgatgaca aaatagcagg 17460
 aagagtgcga aagaaaacaa cgcccatgtt cccttggttac ggctcatatg ccaggcccac 17520
 taatgaaaaa ggaggacagg caacaccgtc cgctagtcaa gacgtgcaaa atcccgaatt 17580
 acaatttttt gccttacta atgtcgccaa tacaccaaaa gcagttctat atgcggagga 17640
 cgtgtcaatt gaagcgccag acactcactt ggtgttcaaa ccaacagtca ctgaaggcat 17700
 tacaagttca gaggctctac tgaccaaca agctgctccc aaccgtccaa actacatagc 17760
 ctttagagat aattttattg gtctcatgta ctacaatagc acaggtaaca tgggagtact 17820
 ggcaggccag gcttctcagc taaatgcagt tggtgacctg caagacagaa atactgagct 17880
 gtcctaccaa ctcatgttg acgccctcgg agaccgcagt cggtaacttt ctatgtggaa 17940
 ccaagctgtg gatagttacg atcctgatgt aagaatcata gaaaaccatg gcgtagaaga 18000
 tgaattgcct aattattgct ttcctttggg aggcattggca gtaaccgaca cctactcgcc 18060
 tataaagggt aatggaggag gcaatggatg ggaagccaat aacggcggtt tcaccgaaag 18120
 aggagtggaa ataggttcag ggaacatgtt tgccatggag attaacctgc aagccaacct 18180
 atggcgtagc tttctgtact ccaatattgg gctgtacctg ccagactctc tcaaaatcac 18240
 tcctgacaac atcacactcc cagagaacaa aaacacctat cagtatatga acggtcgcgt 18300
 gacgccaccc gggctgggtg acacctacgt taacgtgggc gcgcgctggt ccccgatgt 18360
 catggacagt attaacctt ttaatcacca ccgcaacgcc ggactccgct accgttccat 18420
 gctcctggga aacggacgct acgtgccctt ccacatccag gtgcccaga aattctttgc 18480
 aattaaac ctgctgctgc tccccggtt ctacacctac gagtggaaact tccgcaagga 18540
 cgtgaacatg atcttgaga gctcgtggg caatgacctg cgagtggacg gggccagcat 18600
 ccgcttcgac agcatcaacc tgtacgcaa ctttttcccc atggcccaca acacggcctc 18660
 caccctggaa gccatgctgc gcaacgacac caacgaccaa tctttcaacg actacctgtg 18720
 cgcgccaac atgctgtacc ccaccccgcc caacgccacc agcgtgcca tctccattcc 18780
 ctctcgcaac tgggcagcct tcaggggctg gagtttcacc cgctcaaaa ccaaggagac 18840
 cccctcgctg ggctccgggt tcgacccta cttcgtctac tccgggtcca tcccctacct 18900
 ggacggcacc ttctacctca accatacttt caaaaagggtg tcaatcatgt tcgactcctc 18960
 cgtcagctgg cccggcaacg accgtctgct gacgcccac gagttcgaaa tcaagcggtc 19020
 ggtggacggt gaaggggtaca acgtggctca gagcaacatg accaaggact ggttcctgat 19080
 tcagatgctc agccactaca acatcggtc ccagggttc tacgtgcccg aaaattacaa 19140
 ggaccgcatg tactctttct tcagaaactt ccaacccatg agccgcaaaa ttgtagattc 19200

aacggcttac actaattatc aggatgtgaa actgccatac cagcataaca actcaggggtt 19260
 cgtgggctac atgggaccca ccatgcgaga ggggcaggcc taccggcca actatcccta 19320
 tcccctgatt ggggccaccg ccgtgcccag cctcacgcag aaaaagttcc tctgcgaccg 19380
 ggtgatgtgg aggatcccct tctctagcaa cttcatgtct atgggctccc tcaccgacct 19440
 ggggcagaac atgctgtacg ccaactccgc tcacgccttg gatatgacct ttgaggtgga 19500
 tcccatggat gagcccacgc ttctctatgt tctgtttgaa gtcttcgacg tgggtgcgcat 19560
 ccaccagccg caccgcggcg tcacgagggc cgtctacctg cgcacacctt tctctgccgg 19620
 taacgccacc acctaaagaa gccgatgggc tccagcgaac aggagctgca ggccattgtt 19680
 cgcgacctgg gctgcggggc ctactttttg ggcaccttcg acaagcgttt tcccggcttc 19740
 atgtccccc acaagccggc ctgtgccatc gttaacacgg ccggacggga gaccgggggg 19800
 gtccactggc tcgccttcgc ctggaacccg cgtaaccgca cctgctacct gttcgacctt 19860
 tttggtttct ccgacgaaag gctgaagcag atctaccagt tcgagtacga ggggctcctc 19920
 aagcgcagcg ctctggcctc cagccccgac cactgcgta ccctggaaaa gtccacccaa 19980
 acggtccagg ggcccctctc ggccgcctgc gggctcttct gttgcatgtt tttgcacgcc 20040
 ttcgtgcaact ggccctcacac ccccatggat cacaaccca ccatggatct gctcaccgga 20100
 gtgccaaca gcatgcttca cagccccag gtcgccccca ccctgcgccg taaccaggaa 20160
 cacctgtatc gctttctggg gaaacactct gcctattttc gccgccaccg gcagcgcac 20220
 gaacgggcca cggccttcga aagcatgagc caaagagtgt aatcaataaa aaacattttt 20280
 atttgacatg atacgcgctt ctggcgtttt attaaaaatc gaagggttcg agggaggggt 20340
 cctcgtgccc gctggggagg gacacgttgc gatactggaa acgggcgctc caacgaaact 20400
 cggggatcac cagccgcggc aggggcacgt cttctaggtt ctgcttcaa aactgccgca 20460
 ccagctgcag ggctcccatg acgtcggggc ccatatctt gaagtcgcag ttagggccgg 20520
 agtcccgcg gctgttgcg aacacggggg tggcacactg gaacaccagc acgcccgggt 20580
 tgtggatact ggccagggcc gtcgggtcgg tcacctccga cgcattcaga tcctcggcgt 20640
 tgctcagggc aaacgggggt agcttgaca tctgccgcc aatctggggg actaggtcgc 20700
 gcttggtgag gcagtcgcag cgcagaggga tcaggatgcg tcgctgcccg cgttgcatga 20760
 tagggtaact cgccgccagg aactcctcca tttgacggaa ggccatctgg gctttgccgc 20820
 cctcgggtga gaatagcccg caggacttgc tagagaatac gttatgaccg cagttgacgt 20880
 cctccgcgca gcagcggggc tcttcgttct tcagctgaac cacgttgcg ccccaacggg 20940
 tctggaccac cttggctcta gtggggtgct ccttcagcgc ccgctgtccg ttctcgttgg 21000
 ttacatccat ttccaacacg tgctccttgc agaccatctc cactccgtgg aagcaaaaca 21060
 ggacgccctc ctgctgggta ctgcatgct ccatacggc gcatccggtg ggctcccagc 21120
 tcttggtgtt taccgcccg taggcttcca tgtaagccat aaggaatctg cccatcagct 21180
 cggtgagggt cttctggttg gtgaagggtta gcggcaggcc gcggtgctcc tcgttcaacc 21240
 aagtttgaca gatcttgcg tacaccgctc cctggtcggg cagaaactta aaagccgctc 21300

tgctgtcggt	gtctacgtgg	aacttctcca	ttaacatcat	catggtttcc	atacccttct	21360
cccacgctgt	caccagtggg	ttgctgtcgg	ggttcttcac	caacacggcg	gtagaggggc	21420
cctcgccggc	cccgaagtcc	ttcatgggtca	ttctttgaaa	ctccacggag	ccgtccgcgc	21480
gacgtactct	gcgcaccgga	gggtagctga	agcccacctc	caccacgggtg	ccttcgccct	21540
cgctgtcggg	gacaatctcc	ggggatggcg	gcggcgcggg	tgctgccttg	cgagccttct	21600
tcttgggagg	gagctgaggc	gcctcctgct	cgcgctcggg	gctcatctcc	cgcaagtagg	21660
gggtaatgga	gctgcctgct	tggttctgac	ggttggccat	tgtatcctag	gcagaaagac	21720
atggagctta	tgcgcgagga	aactttaacc	gccccgtccc	ccgtcagcga	cgaagatgtc	21780
atcgctgaac	aggacccggg	ctacgttacg	ccgcccggag	atctggaggg	gcctgaccgg	21840
cgcgacgcta	gtgagcggca	ggaaaatgag	aaagaggagg	cctgctacct	cctggaaggc	21900
gacgttttgc	taaagcattt	cgccaggcag	agcaccatag	ttaaggaggc	cttgcaagac	21960
cgctccgagg	tgcccttgga	cgctcgccgc	ctctcccagg	cctacgaggc	gaaccttttc	22020
tcgcctcgag	tgccctcgaa	gagacagccc	aacggcacct	gcgagcccaa	cccgcgactc	22080
aacttctacc	ccgtgttcgc	cgtaccagag	gcgctggcca	cctatcacat	ttttttcaaa	22140
aaccaacgca	tccccctatc	gtgccggggc	aaccgcaccg	cggccgatag	gaatctcagg	22200
cttaaaaacg	gagccaacat	acctgatatc	acgtcgctgg	aggaagtgcc	caagattttc	22260
gaggggtctg	gtcgagatga	gaagcggggc	gcgaacgctc	tgagaaaaga	acagaaaagag	22320
agtcagaacg	tgctgggtgga	gctggagggg	gacaacgcgc	gtctggccgt	cctcaaacgc	22380
tgcatagaag	tctccactt	cgcctacccc	gccctcaact	tgccacccaa	agttatgaaa	22440
tcgggtcatg	atcagctgct	catcaagaga	gctgagcccc	tggtatcccga	ccaccccgag	22500
gcggaaaaact	cagaggacgg	aaagcccgtc	gtcagcgacg	aggagctcga	gcgggtggctg	22560
gaaaccaggg	acccccaaca	gttgcaagag	aggcgcaaga	tgatgatggc	ggccgtgctg	22620
gtcaccgtgg	agctggaatg	cctgcaacgg	tttttcagcg	acgtggagac	gctacgcaaa	22680
atcgggggaat	ccctgcacta	caccttcgcg	cagggtctacg	tccgccaggc	ctgcaagatc	22740
tccaacgtgg	agctcagcaa	cctgggtctc	tacatgggca	tcctccacga	gaaccggctg	22800
gggcagagcg	tgctgcactg	caccttgcaa	ggcgaggcg	ggcgggacta	cgtgcgagac	22860
tgcatctacc	tcttcctcac	cctcacctgg	cagaccgcca	tgggcgtctg	gcagcagtgc	22920
ttggaagaga	gaaacctcaa	agagctagac	aaactcctct	gccgccagcg	gcgcgccttg	22980
tggtccgggt	tcagcgagcg	cacggctcgc	agcgctctgg	cggacatcat	cttcccggag	23040
cgcctgatga	aaaccttgca	aaacggcctg	ccggatttca	tcagtcaaag	cattttgcaa	23100
aacttccgct	cttttgtcct	ggaacgctcc	gggatcttgc	ccgccatgag	ctgcgcgcta	23160
ccttctgact	ttgtccccct	ctcctaccgc	gagtgccttc	ccccactgtg	gagccactgc	23220
tacctcttcc	aactggccaa	ctttctggcc	taccactccg	acctcatgga	agacgtaagc	23280
ggagaggggt	tactggagtg	ccactgccgc	tgcaacctgt	gcacccccca	cagatcgctg	23340

gcctgcaaca ccgagctact cagcgaaacc cagggtcatag gtaccttcga gatccagggg 23400
 ccccgagcagc aagaggggtgc ttccggcttg aagctcactc cggcgctgtg gacctcggct 23460
 tacttacgca aatttgtagc cgaggactac cacgcccaca aaattcagtt ttacgaagac 23520
 caatctcgac caccgaaagc cccctcacg gcctgctca tcaccagag caagatcctg 23580
 gccaattgc aatccatcaa ccaagcgcgc cgcgatttcc ttttgaaaaa gggtcggggg 23640
 gtgtacctgg acccccagac cggcgaggaa ctcaaccgt ccacactctc cgtcgaagca 23700
 gccccccga gacatgccgc ccaaggaac cgccaagcag ctgatcgtc ggcagagagc 23760
 gaagaagcaa gagctgctcc agcagcaggt ggaggacgag gaagagatgt gggacagcca 23820
 ggcagaggag gtgtcagagg acgaggagga gatggaaagc tgggacagcc tagacgagga 23880
 ggaggacgag ctttcagagg aagaggcgac cgaagaaaaa ccacctgcat ccagcgcgcc 23940
 ttctctgagc cgacagccga agccccggcc ccgacgccc ccggccggct cactcaaagc 24000
 cagccgtagg tgggacgcca ccgaatctcc agcggcagcg gcaacggcag cgggtaaggc 24060
 caaacgcgag cggcggggggt attgctcctg gcgggcccac aaaagcagta ttgtgaactg 24120
 cttgcaacac tgcgggggaa acatctcctt tgcccgcgc tacctcctct tccatcacgg 24180
 tgtggccttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24240
 gctcggagaa aaaagctaag gcctcctccg ccgcgaggaa aaactccgcc gccgctgccg 24300
 ccgccaagga tccaccggcc accgaagagc tgagaaaagc catctttccc actctgtatg 24360
 ctatctttca gcaaagccgc gggcagcacc ctacgcgca actgaaaata aaaaaccgct 24420
 ccttccgctc gctcaccgc agctgtctgt accacaagag agaagaccag ctgcagcgca 24480
 ccctggacga cgccgaagca ctgttcagca aatactgctc agcgtctctt aaagactaaa 24540
 agaccgcgc tttttcccc tcggccgcca aaaccacgt catcgccagc atgagcaagg 24600
 agattcccac ccctacatg tggagctatc agcccagat gggcctggcc gcgggggccc 24660
 ccaggaacta ctccagcaag atgaactggc tcagcgccgg cccccacatg atctcacgag 24720
 ttaacggcat ccgagcccac cgaaaccaga ttctcttaga acaggcgga atcaccgcca 24780
 caccgcggcg ccaactcaac ccgcctagtt ggcccgcgc ccagggtgat caggaaaatc 24840
 cccgcccgc cacagtcctc ctgccacgc acgcggaggc cgaagtcctc atgactaact 24900
 ctggggtaca attagcgggc ggggtccaggt acgccaggta cagaggtcgg gccgctcctt 24960
 actctcccgg gagtataaag agggatgatca ttcgaggccg aggtatccag ctcaacgacg 25020
 agacgggtgag ctctcaacc ggtctcagac ctgacggagt cttccagctc ggaggagcgg 25080
 gccgctcttc cttcaccact cgccaggcct acctgaccct gcagagctct tcctcgacg 25140
 cgcgctccgg gggaatcggc actctccagt tctgtggaaga gttcgttccc tccgtctact 25200
 tcaaccctt ctccggctcg cctggacgct acccgacgc cttcattccc aactttgacg 25260
 cagtgagtga atccgtggac ggctacgact gatgacagat ggtgcggccg tgagagctcg 25320
 gctgcgacat ctgcatcact gccgtcagcc tcgctgctac gtcgggagg cgatcgtctt 25380
 cagctacttt gagctgccgg acgagcacc tcagggtccg gctcacgggt tgaaactcga 25440

gatcgagaac	gcgctcgagt	ctcgcctcat	cgacaccttc	accgcccgac	ctctcctggt	25500
agaaatccaa	cgggggatca	ctaccatcac	cctgtttctgc	atctgccccca	cgcccggatt	25560
acatgaagat	ctgtgttggtc	atctttgctgc	tcagtttaaat	aaaaactgaa	ctttttgccg	25620
caccttcaac	gccatctgtg	atttctacaa	caaaaagtgc	ttctggcaaa	ggtacacaaa	25680
ctgtattttta	ttctaattct	acctcatcta	tcgtgctgaa	ctgcgcctgc	actaacgaac	25740
ttatccagtg	gattgcaaac	ggtagtgtgt	gcaagtactt	ttgggggaac	gatatagtta	25800
gtagaaataa	cagcctttgc	gagcactgca	actcctccac	actaatcctt	tatcccccat	25860
ttgttactgg	atggtatatg	tgcgttggct	ccggtttaaa	tcctagtgtc	tttcataagt	25920
ggtttctaca	aaaagagacc	cttcccaaca	attctgtttc	ttttttcgcc	ctatcctact	25980
gctgtttctcc	ctctgggttac	tctttcaaac	ctctaattgg	tatttttagct	ttgatactca	26040
taatctttat	taactttata	ataattaaca	acttacagta	aacatgcttg	ttctactgct	26100
cgccacatct	ttcgctctct	ctcacgccag	aacaagtatt	gttggcgag	gttacaatgc	26160
aactcttcaa	tctgcttaca	tgccagattc	cgaccagata	ccccatatta	cgtgggtactt	26220
acaaacctcc	aaacctaatt	cttcatttta	tgaaggaaac	aaactctgag	atgactccga	26280
caacagaacg	cacacatttc	cccacccttc	actacaattc	gaatgcgtaa	acaaaagctt	26340
gaagctttac	aacttaaagc	cttcagattc	tggcttgtag	catgctgtag	ttgaaaaaag	26400
taatttagaa	gtccacagtg	attacattga	attgacgggt	gtggacctgc	cacctccaaa	26460
atgtgagggt	tcctcctctt	accttgaagt	tcaaggcgtg	gatgcctact	gcctcatata	26520
cattaactgc	agcaactcta	aatatccagc	tagaattttac	tataatggac	aggaaagtaa	26580
tcttttttat	tatttaacaa	caagcgtggt	taacggtaaa	cagttacctg	actatttttac	26640
tgtgtgtgtt	gaattttcca	cctacagaga	aacgtatgcc	aagcggcctt	acaattttctc	26700
atacccggtt	aacgaccttt	gcaatgaaat	acaagcgtc	gaaactggaa	ctgatttttac	26760
tccaattttc	attgctgcca	ttgttgtaag	cttaattacc	attattgtca	gcctagcatt	26820
ttactgcttt	tacaagccca	aaaaccctaa	gtttgaaaaa	cttaaaactaa	aacctgtcat	26880
tcaacaagtg	tgattttgtt	ttccagcatg	gtagctgcat	ttctacttct	cctctgtcta	26940
cccatcattt	tcgtctcttc	aactttcgcc	gcagtttccc	acctggaacc	agagtgccta	27000
ccgccttttg	acgtgtatct	gattctcacc	tttgtttgtt	gtatatccat	ttgcagtata	27060
gcctgctttt	ttataacaat	ctttcaagcc	gccgactatt	tttacgtgag	aattgcttac	27120
tttagacacc	atcctgaata	cagaaatcaa	aacgttgcct	ccttactttg	tttggcatga	27180
ttaagtattt	gctgatactt	aattattttac	ccctaataca	ctgtaattgt	ccattcacca	27240
aaccttggtc	attctacacc	tgttatgata	aaatccccga	cactcctggt	gcttggcttt	27300
acgcagccac	cgccgctttg	gtattttatat	ctacttgcct	tggagtaaaa	ttgtattttta	27360
ttttacacac	tgggtggcta	catcccagag	aagatttacc	tagatatacct	cttgtaaacg	27420
cttttcaatt	acagcctctg	cctcctcctg	atcttcttcc	tcgagctccc	tctattgtga	27480

gctacttttca actcaccggt ggagatgact gactctcagg acattaatat tagtgtggaa 27540
agaatagctg ctcagcgtca gcgagaaacg cgagtgttg aatacctgga actacagcaa 27600
cttaaagagt cccactggtg tgagaaaagga gtgctgtgcc atgttaagca ggcagccctt 27660
tcctacgatg tcagcgttca gggacatgaa ctgtcttaca ctttgccttt gcagaaacaa 27720
accttctgca ccatgatggg ctctacctcc atcacaatca cccaacaagc cgggcctgta 27780
gagggggcta tcctctgtca ctgtcacgca cctgattgca tgtccaaact aatcaaaact 27840
ctctgtgctt taggtgatat ttttaagggtg taaatcaata ataaacttac cttaaatttg 27900
acaacaaatt tctggtgaca tcattcagca gcaccacttt accctcttcc cagctctcgt 27960
atgggatgcg atagtgggtg gcaaacttcc tccaaaccct aaaagaaata ttggtatcca 28020
cttccttgtc ctcaccaca attttcatct tttcatagat gaaaagaacc agagttgatg 28080
aagacttcaa ccccgcttac ccctatgaca ccacaaccac tcctgcagtt ccctttatat 28140
cacccttctt tgtaaacagc gatggctctc aggaaaaccc cccaggtggt ttaagtctgc 28200
gaatagctaa acccctatat ttcgacatgg agagaaaact agccctttca cttggaagag 28260
ggttgacaat taccgccgcc ggacaattag aaagtacgca gagcgtacaa accaaccac 28320
cgttgataat taccaacaac aacacactga ccctacgtca ttctcccccc ttaaaccctaa 28380
ctgacaatag cttagtgtca ggctactcga gtcctctccg cgtcacagac acaaaactta 28440
catttaactt cacatcacca ctccgttatg aaaatgaaaa cttactttt aactatacag 28500
agcctcttaa acttataaat aacagccttg ccattgacat caattcctca aaaggcctta 28560
gtagcgtcgg aggctcacta gctgtaaacc tgagttcaga cttaaagttt gacagcaacg 28620
gatccatagc ttttggcata caaacctgt ggaccgctcc gacctcgact ggcaactgca 28680
ccgtctacag cgagggcgat tccctactta gtctctgttt aaccaaatgc ggagctcacg 28740
tcttaggaag tgtaagttta accggtttta caggaaccat aaccctaatg actgatattt 28800
ctgtcaccat tcaatttaca tttgacaaca atggtaagct actaagctct ccacttataa 28860
acaacgcctt tagtattcga cagaatgaca gtacggcctc aaaccctacc tacaacgccc 28920
tggcgtttat gcctaacagt accatatatg caagaggggg aggtggtgaa ccacgaaaca 28980
actactacgt ccaaacgtat cttaggggaa atgttcaaaa accaatcatt cttactgtaa 29040
cctacaactc agtcgccaca ggatattcct tatcttttaa gtggactgct cttgcacgtg 29100
aaaagtttgc aacccaaca acctcgtttt gctacattac agaacaataa aaccgtgtac 29160
cccaccgttt cgtttttttc agatgaaacg ggcgagagtt gatgaagact tcaaccagt 29220
gtacccttat gacccccac atgtcctgt tatgcccttc attactccac cttttacctc 29280
ctcggatggg ttgcaggaaa aaccacttgg agtggttaagt ttaaaactaca gagatcccat 29340
tactacgcaa aatgagtctc ttacaattaa actaggaaac ggcctcactc tagacaacca 29400
gggacaacta acatcaaccg ctggcgaagt agaacctca ctactaacg ctaacaacaa 29460
acttgcactg gtctatagcg atcctttagc agtaaagcgc aacagcctaa ccttatcgca 29520
caccgctccc cttgttattg ctgataactc ttttagcattg caagtttcag agcctatttt 29580

tataaatgac	aaggacaaac	tagccctgca	aacagccgcg	ccccttgtaa	ctaacgctgg	29640
cacccttcgc	ttacaaagcg	ccgccccttt	aggcattgca	gaccaaacc	taaaactcct	29700
gtttaccaac	cctttgtact	tgcagaataa	ctttctcacg	ttagccattg	aacgaccctt	29760
tgccattacc	aatactggaa	agctggctct	acagctctcc	ccaccgctac	aaacagcaga	29820
cacaggcttg	actttgcaaa	ccaacgtgcc	attaactgta	agcaacggga	ccctaggctt	29880
agccataaag	cgcccactta	ttattcagga	caacaacttg	tttttggaact	tcagagctcc	29940
cctgcgtctt	ttcaacagcg	accagtgact	agggcttaac	ttttacaccc	ctcttgcggt	30000
acgcgatgag	gcgctcactg	ttaacacagg	ccgcggcctc	acagtgagtt	acgatgggtt	30060
aattttaaat	cttggttaagg	atcttcgctt	tgacaacaac	accgtttctg	tcgctcttag	30120
tgctgctttg	cctttacaat	acactgatca	gcttcgcctt	aacgtgggcg	ctgggctgcg	30180
ttacaatcca	gtgagtaaga	aattggacgt	gaacccaat	caaaacaagg	gtttaacctg	30240
ggaaaatgac	tacctcattg	taaagctagg	aaatggatta	ggttttgatg	gcgatggaaa	30300
catagctgtt	tctcctcaag	ttacatcgcc	tgacacctta	tggaccactg	ccgaccctac	30360
ccccaatgtt	tccatctaca	ctgatttaga	tgccaaaatg	tggctctcgt	tggtaaaaca	30420
agggggtgtg	gttcacgggt	ctgttgcttt	aaaagcattg	aaaggaacc	tattgagtcc	30480
tacggaaaagc	gccattgtta	ttatactaca	ttttgacaat	tatggagtgc	gaattctcaa	30540
ttatcccact	ttgggcactc	aaggcacgtt	gggaaataat	gcaacttggg	gttataggca	30600
gggagaatct	gcagacacta	atgtactcaa	tgcactagca	tttatgccca	gttcaaaaag	30660
gtaccaaga	gggcgtggaa	gcgaagttca	gaatcaaaact	gtgggctaca	cttgatataca	30720
gggtgacttt	tctatgccc	taccgtacca	aatacagtac	aactatggac	caactggcta	30780
ctcctttaaa	tttatttgga	gaactgtttc	aagacaacca	tttgacatcc	catgctgttt	30840
tttctcttac	attacggaag	aataaaacaa	ctttttcttt	ttattttctt	tttattttac	30900
acgcacagta	aggcttcctc	cacccttcca	tctcacagca	tacaccagcc	tctccccctt	30960
catggcagta	aactgttggtg	agtcagtccg	gtatttgagg	gttaagatcc	aaacagtctc	31020
tttggtgatg	aaacatggat	ccgtgatgga	cacaaatccc	tgggacaggt	tctccaacgt	31080
ttcggtaaaa	aactgcatgc	cgccctacaa	aacaaacagg	ttcaggctct	ccacgggtta	31140
tctccccgat	caaactcaga	cagagtaaag	gtgcgatgat	gttcactaa	accacgcagg	31200
tggcgctgtc	tgaacctctc	ggtgcgactc	ctgtgaggct	ggtaagaagt	tagattgtcc	31260
agcagcctca	cagcatggat	catcagtcta	cgagtgcgtc	tggcgagca	gcgcatctga	31320
atctcactga	gattccggca	agaatcgac	accatcacaa	tcagggttgtt	catgatccca	31380
tagctgaaca	cgctccagcc	aaagctcatt	cgctccaaca	gcgccaccgc	gtgtccgtcc	31440
aaccttactt	taacataaat	cagggtgtctg	ccgcgtacaa	acatgctacc	cgcatacaga	31500
acctcccggg	gcaaaccctt	gttcaccacc	tgctgttacc	agggaacct	cacatttatc	31560
agggagccat	agatagccat	tttaaaccaa	ttagctaaca	ccgccccacc	agctctacac	31620

tgaagagaac cgggagagtt acaatgacag tgaataatcc atctctcata acccctaattg 31680
gtctgatgga aatccagatc taacgtggca cagcagatac acactttcat atacattttc 31740
atcacatgtt tttcccaggc cgtaaaata caatcccaat acacgggcca ctctgcagt 31800
acaataaagc taatacaaga tggatatactc ctcacctcac taacattgtg catgttcata 31860
ttttcacatt ctaagtaccg agagtctctc tctacaacag cactgccgcg gtcctcacia 31920
ggtaggtggt ggtgacgatt gtaaggagcc agtctgcagc gataccgtct gtcgcgttgc 31980
atcgtagacc agggaccgac gcaacttctc gtacttgtag tagcagaacc acgtccgctg 32040
ccagcacgtc tccaagtaac gccggtccct gcgtcgctca cgctccctcc tcaacgcaaa 32100
gtgcaaccac tcttgtaatc cacacagatc cctctcggcc tccggggcga tgcacacctc 32160
aaacctacag atgtctcggc acagtccaac acacgtagtg agggcgagtt ccaaccaaga 32220
cagacagcct gatctatccc gacacactgg aggtggagga agacacggaa gaggcatgtt 32280
attccaagcg attcaccacg gggtcgaaat gaagatcccg aagatgacaa cggtcgcctc 32340
cggagccctg atggaattta acagccagat caaacattat gcgattttcc aggtatcaaa 32400
tcgcgccctc caaaagagcc tggaccgcga cttccacaaa caccagcaaa gcaaaagcgt 32460
tattatcaaa ctcttcgatc atcaagctgc aggtactgtac aatgccaag taattttcat 32520
ttctccactc gcgaatgatg tcgcggaacaa tagtctgaag gttcatgccg tgcataataa 32580
aaagctccga aaggcgccc tctatagcca tgcgtagaca caccatcatg actgcaagat 32640
atcgggctcc tgagacacct gcagcagatt taacagaccc aggtcaggtt gctctccgcg 32700
atcgcgatc tccatccgca aagtcatttg caaataatta aatagatctg cgccgactaa 32760
atctgttaac tccgcgctag gaactaaatc aggtgtggct acgcagcaca aaagttccag 32820
ggatggcgcc aaactcacta gaaccgctcc cgagtagcaa aactgatgaa tgggagtaac 32880
acagtgtaaa atgttcagcc aaaaatcact aagctgctcc tttaaaaagt ccagtacttc 32940
tatattcagt tcgtgcaagt actgaagcaa ctgtgcggga atatgcacag caaaaaaat 33000
agggcggtc agatacatgt tgacctaaaa taaaaagaat cattaaacta aagaagcctg 33060
gcgaacggtg ggatatatga cagctccag cagcaggcaa gcaaccggct gtccccggga 33120
accgcggtaa aattcatccg aatgattaaa aagaacaaca gagacttccc accatgtact 33180
cggttgatc tctgagcac agagcaatac cccctcaca ttcataatccg ctacagaaaa 33240
aaaacgtccc agataccag cgggaatatc caacgacagc tgcaaagaca gcaaaacaat 33300
ccctctggga gcaatcacia aatcctccgg tgaaaaaagc acatacatat tagaataacc 33360
ctgttgctgg ggcaaaaagg cccgtcgtcc cagcaaatgc acataaatat gttcatcagc 33420
cattgccccg tcttaccgcg taaacagcca cgaaaaaatc gagctaaaat ccaccaaca 33480
gcctatagct atatatacac tccaccaat gacgctaata ccgcaccacc cacgacaaa 33540
gttcaccac acccaciaaa cccgcgaaaa tccagcgccg tcagcacttc cgcaatttca 33600
gtctcaciaa gtcacttccg cgcgcctttt cactttcca cacacgccct tcgcccgc 33660
gccctcgcg caccgcgct caccacgt caccgcagc caccgcggcc ccgcctcgt 33720

```

cctccccgct cattatcata ttggcacgtt tccagaataa ggtatattat tgatgcagca 33780
aaacaatccc tctgggagca atcacaaaat cctccggtga aaaaagcaca tacatattag 33840
aataaccctg ttgctggggc aaaaaggccc gtcgtcccag caaatgcaca taaatatgtt 33900
catcagccat tgccccgtct taccgcgtaa acagccacga aaaaatcgag ctaaaatcca 33960
cccaacagcc tatagctata tatacactcc acccaatgac gctaataccg caccacccac 34020
gaccaaagtt caccacacacc caaaaaacc gcgaaaatcc agcgccgtca gcacttccgc 34080
aatttcagtc tcacaacgtc acttccgcgc gccttttcac tttccacac acgcccttcg 34140
cccgcccgcc ctgcgcgccac ccgcggtcac ccacggtcac cgcacgtcac ccgggccccg 34200
cctcgtcctt ccccgctcat tatcatattg gcacgtttcc agaataagggt atattattga 34260
tgca 34264

```

```

<210> 5
<211> 31044
<212> DNA
<213> simian adenovirus SV-25

```

```

<400> 5
catcatcaat aatatacctt attctggaaa cgtgccata tgataatgag cggggaggag 60
cgaggcgggg ccggggtgac gtgcggtgac gcggggtggc gcgaggcgcg ggcgaagggc 120
gcgggtgtgt gtgtgggagg cgcttagttt ttacgtatgc ggaaggagggt ttatataccgg 180
aagatgggta atttgggcgt atacttgtaa gttttgtgta atttggcgcg aaaactgggt 240
aatgaggaag ttgagggttaa tatgtacttt ttatgactgg gcggaatttc tgctgatcag 300
cagtgaactt tgggcgctga cggggagggt tcgctacgtg acagtaccac gagaaggctc 360
aaagggtcca ttattgttac tcttcagcgt ttcgctggg tatttaaacg ctgtcagatc 420
atcaagaggc cactcttgag tgctggcgag aagagttttc tcctccgtgc tgccacgatg 480
aggctggtcc ccgagatgta cgggtgtttt agcgacgaga cgggtgcgtaa ctcatgatgac 540
ctgctgaatt cagacgcgt ggaaatttcc aattcgctg tgctttcgcc gccgtcactt 600
cacgacctgt ttgtgttttg gctcaacgct tagcaacgtg ttatataggg tcaagaagga 660
gcaggagacg cagtttgcta ggctgttggc cgatactcct ggagtttttg tggctctgga 720
tctaggccat cactctcttt tccaagagaa aattatcaa aacttaactt ttacgtctcc 780
tggtcgacg gttgcttccg ctgcctttat tacctatatt ttggatcaat ggagcaacag 840
cgacagccac ctgtcgtggg agtacatgct ggattacatg tcgatggcg tggtggagggc 900
catgctgcgg aggaggggtt gcatttactt gcgggcgag cctccgcggc tggaccgagt 960
ggaggaggag gacgagccgg gggagaccga gaacctgagg gccgggctgg accctccaac 1020
ggaggactag gtgctgagga tgatcccga gaggggacta gtggggctag gaagaagcaa 1080
aagactgagt ctgaacctcg aaactttttg aatgagttga ctgtgagttt gatgaatcgt 1140
cagcgtccgg agacaatttt ctggtctgaa ttggaggagg aattcaggag gggggaactg 1200
aacctgctat acaagtatgg gtttgaacag ttaaaaactc actggttgga gccgtgggag 1260

```

gattttgaaa ccgccttgga cacttttgct aaagtggctc tgcggccgga taaggtttac	1320
actatccgcc gcactgttaa cataaagaag agtgtttatg ttataggcca tggagctctg	1380
gtgcagggtgc aaaccgtcga ccgggtggcc tttagttgcg gtatgcaaaa tctgggcccc	1440
ggggtgatag gcttaaattg tgtaacattt cacaatgtaa ggtttactgg tgaaagtttt	1500
aacggctctg tgtttgcaaa taacacacag ctgacgctcc acggcgttta cttttttaac	1560
tttaataaca catgtgtgga gtcgtggggc aggggtgtctt tgaggggctg ctgttttcac	1620
ggctgctgga aggcgggtggg ggggaagactt aaaagtgtaa catctgtaaa aaaatgcgtg	1680
tttgagcggg gtgtgttggc tttaactgtg gagggctgtg gacgcattag gaataatgcg	1740
gcgtctgaga atggatgttt tcttttgcta aaaggcacgg ctagtattaa gcataacatg	1800
atatgcggca gcgggtctgta cccttcacag ctgttaactt gcgcggatgg aaactgtcag	1860
accttgcgca ccgtgcacat agcgtccac cagcgccgcg cctggccaac attcgagcac	1920
aatatgctta tgcgttgtgc cgtccacttg ggccctaggc gaggcgtgtt tgtgccttac	1980
cagtgttaact ttagccatac caagatttta ctagaacctg ataccttctc tcgagtgtgt	2040
ttcaatgggg tgtttgacat gtcaatggaa ctgtttaaag tgataagata tgatgaatcc	2100
aagtctcggt gtcgcccattg tgaatgcgga gctaatacctc tgagggtgta tcctgtaacc	2160
ctaaacgtta ccgaggagct gaggacggat caccacatgt tgtcctgcct gcgcaccgac	2220
tatgaatcca gcgacgagga gtgagggtgag gggcgagacc acaaagggtg taaaggggag	2280
tgaggggtgg gtgtgatgat tcaaaatgag cgggacgacg gacggcaacg cgtttgaggg	2340
tggagtgttc agcccttatc tgacatctcg tcttccttcc tgggcaggag tgcgtcagaa	2400
tgtagtgggc tccaccgtgg acggacgacc ggtcgcccct gcaaattccg ccaccctcac	2460
ctatgccacc gtgggatcat cgttggacac tgccgcggca gctgccgctt ctgctgccgc	2520
ttctactgct cgcggcatgg cggctgattt tggactgtat aaccaactgg cactgcagc	2580
tgtggcgtct cggtctctgg ttcaagaaga tgccctgaat gtgatcctga ctgcctgga	2640
gatcatgtca cgtcgcttgg acgaactggc tgcgcagata tcccaagcta accccgatac	2700
cacttcagaa tcctaaaata aagacaaaca aatatgttga aaagtaaaat ggctttattt	2760
gttttttttg gctcggtagg ctggtgtcca cctgtctcgg tcgttaagaa ctttgtgtat	2820
gttttccaaa acacgggtaca gatgggcttg gatgttcaag tacatgggca tgaggccatc	2880
tttgggggtga agataggacc attgaagagc gtcattgctc ggggtggtgt tgtaaattac	2940
ccagtctgtag cagggtttct gggcgtggaa ctggaagatg tccttttagga gtaggctgat	3000
ggccaagggc aggcccttag ttaggtgtt tacaagcg ttaagctggg agggatgcat	3060
gcgggggggag atgatatgca tcttggttg gatcttgagg ttagctatgt taccaccag	3120
gtctctgcgg gggttcatgt tatgaaggac caccagcacg gtgtagccgg tgcatttggg	3180
gaacttgtca tgcagtttg aggggaaggc gtggaagaat ttagagaccc cttgtggcc	3240
ccctagggtt tccatgcact catccataat gatggcaatg ggaccctgg cgccgcttt	3300

ggcaaacacg ttttgggggt tggaaacatc atagttttgc tctagagtga gctcatcata 3360
ggccatctta acaaagcggg gtaggagggg gcccgactgg gggatgatag ttccatctgg 3420
gcctggggcg tagttaccct cacagatctg catctcccag gccttaattt ccgagggggg 3480
tatcatgtcc acctgggggg caataaagaa cacggtttct ggcgggggat tgatgagctg 3540
ggtggaaagc aagttacgca gcagttgaga tttgccacag ccggtggggc cgtagatgac 3600
cccgatgacg ggttgacgct ggtagttgag agaggaacag ctgccgtcgg ggcgcaggag 3660
gggggctacc tcattcatca tgcttctaac atgtttatit tcaactacta agttttgcaa 3720
gagcctctcc ccaccaggg ataagagttc ttccaggctg ttgaagtgtt tcagcggttt 3780
taggccgtcg gccatgggca tcttttcgag cgactgacga agcaagtaca gtcggtccca 3840
gagctcggtg acgtgctcta tggaatctcg atccagcaga cttcttggtt gcggggggtg 3900
ggtcgacttt cgctgtaggg caccagccgg tgggcgtcca gggccgcgag ggttctgtcc 3960
ttccagggtc tcagcgtccg ggtgaggggt gtctcggtga cggtgaaggg atgagccccg 4020
ggctggggcg ttgcgagggg gcgcttcagg ctcatcctgc tgggtgctgaa gcggacgtcg 4080
tctccctgtg agtcggccag atagcaacga agcatgaggt cgtagctgag ggactcggcc 4140
gcgtgtccct tggcgcgag ctttcccttg gaaacgtgct gacatttggt gcagtgcaga 4200
cattggaggg cgtagagttt gggggccagg aagaccgact cgggcgagta ggcgtcggct 4260
ccgcactgag cgcagacggt ctgcactcc actagccacg tgagctcggg tttagcggga 4320
tcaaaaacca agttgcctcc attttttttg atgcgtttct taccttgctt ttccatgagt 4380
ttgtggcccc cttccgtgac aaaaaggctg tcggtgtctc cgtagacaga cttgaggggg 4440
cgatcttcca aagggtgttc gaggtcttcc gcgtacagga actgggacca ctccgagacg 4500
aaggctctgg tccaggctaa cacgaaggag gcaatctgcg aggggtatct gtcgttttca 4560
atgagggggg ccaccttttc cagggtgtgc agacacaggt cgtcctcctc cgcgtccacg 4620
aagggtgattg gcttgtaagt gtaggtcacg tgatctgcac ccccaaagg ggtataaaag 4680
ggggcggtgc caccctctcc gtcactttct tccgcatcgc tgtggaccag agccagctgt 4740
tcgggtgagt aggccctctc aaaagccggc atgatctcgg cgctcaagtt gtcagtttct 4800
acaaacgagg tggatttgat attcacgtgc cccgcggcga tgcttttgat ggtggagggg 4860
tccatctgat cagaaaacac gatctttttg ttgtcaagtt tgggtggcgaa agacccgtag 4920
agggcgttgg aaagcaactt ggcgatggag cgcaggggtc gattttttct ccgatcggcc 4980
ctctccttgg cggcgatggt gagttgcacg tactcccggg ccgcgcaccg cactcgggg 5040
aacacggcgg tgcgtcgtc gggcaggatg cgcacgcgcc agccgcgatt gtgcaggggtg 5100
atgaggtcca cgctggtagc cacctccccg cggaggggct cgttgggtcca acacaatcgc 5160
cccccttttc tggagcagaa cggaggcagg ggatctagca agttggcggg cggggggctg 5220
gcgtcgatgg tgaagatacc gggtagcagg atcttattaa aataatcgat ttcgggtgtcc 5280
gtgtcttgca acgcgtcttc ccacttcttc accgccaggg ccctttcgta gggattcagg 5340
ggcgggtccc agggcatggg gtgggtcagg gccgaggcgt acatgccgca gatgtcatac 5400

acgtacaggg	gttccctcaa	cacccccgatg	taagtgggggt	aacagcgccc	cccgcggatg	5460
ctggctcgca	cgtagtcgta	catctcgcg	gagggagcca	tgaggccgtc	tcccaagtgg	5520
gtcttgtggg	gtttttcggc	ccggtagagg	atctgtctga	agatggcgtg	ggagttggaa	5580
gagatggtgg	ggcgttggaa	gacgttaaag	ttggccccgg	gtagtccac	ggagtcttgg	5640
atgaactggg	cgtaggattc	ccggagtttg	tccaccaggg	cggcggtcac	cagcacgtcg	5700
agagcgcagt	agtccaacgt	ctcgcggaac	aggttgtagg	ccgtctcttg	ttttttctcc	5760
cacagttcgc	ggttgaggag	gtattcctcg	cggtctttcc	agtactcttc	ggcgggaaat	5820
cctttttcgt	ccgctcggt	agaacctaac	atgtaaaatt	cgttcaccgc	tttgatgga	5880
caacagcctt	tttctaccgg	cagggcgtag	gcttgagcgg	cctttctgag	agaggtgtgg	5940
gtgagggcga	aggtgtccc	caccatcact	ttcaggtact	gatgtttgaa	gtccgtgtcg	6000
tcgcaggcgc	cctgttccca	cagcgtgaag	tcggtgcgct	ttttctgcct	gggattgggg	6060
agggcgaaag	tgacatcggt	aaagagtatt	ttcccggcgc	ggggcatgaa	gttgcgagag	6120
atcctgaagg	gcccgggcac	gtccgagcgg	ttgttgatga	cctgcgccgc	caggacgatc	6180
tcgtcgaagc	cgttgatgtt	gtgaccacg	atgtaaagt	cgatgaagcg	cggctgtccc	6240
ttgagggccg	gcgctttttt	caactcctcg	taggtgagac	agtcggcgga	ggagagaccc	6300
agctcagccc	gggcccagtc	ggagagtga	ggattagccg	caaggaagga	gctccataga	6360
tccaaggcca	ggagagtttg	caagcggtcg	cggaaactcg	ggaacttttt	ccccacggcc	6420
attttctccg	gtgtcactac	gtaaaagggt	ttggggcggt	tgttccacac	gtcccatcgg	6480
agctctaggg	ccagctcgca	ggcttggcga	acgaggggtct	cctcgccaga	gacgtgcatg	6540
accagcataa	agggtagcaa	ctgtttcccg	aacgagccca	tccatgtgta	ggtttctacg	6600
tcgtaggtga	caaagagccg	ctgggtgcgc	gcgtgggagc	cgatcgga	gaagctgatc	6660
tcctgccacc	agctggagga	atgggtgtta	atgtgggtga	agtagaagtc	ccgccggcgc	6720
acagagcatt	cggtctgatg	tttgtaaaag	cgaccgcagt	agtcgcagcg	ctgcacgctc	6780
tgtatctcct	gaacgagatg	cgcttttcgc	ccgcgcacca	gaaaccggag	ggggaagtgt	6840
agacgggggg	ctggtggggc	gacatcccct	tcgccttggc	ggtgggagtc	tgctctgctg	6900
tcctccttct	ctgggtggac	gacgggtggg	acgacgacgc	cccgggtgcc	gcaagtccag	6960
atctccgcca	cggaggggtg	caggcgctgc	aggaggggac	gcagctgccc	gctgtccagg	7020
gagtcgaggg	aagtcgcgct	gaggtcggcg	ggaagcgttt	gcaagttcac	tttcagaaga	7080
ccggtaagag	cgtgagccag	gtgcagatgg	tacttgattt	ccaggggggt	gttgatgaa	7140
gcgtccacgg	cgtagaggag	tccgtgtccg	cgcggggcca	ccaccgtgcc	ccgaggaggt	7200
tttatctcac	tcgtcgaggg	cgagcgccgg	ggggtagagg	cggctctgctg	ccggggggca	7260
gcggaggcag	aggcacgttt	tcgtgaggat	tcggcagcgg	ttgatgacga	gcccggagac	7320
tgctggcgtg	ggcgacgacg	cggcggttga	ggctctggat	gtgccgtctc	tgctgaaga	7380
ccaccggccc	ccgggtcctg	aacctaaaga	gagttccaca	gaatcaatgt	ctgcatcggt	7440

aacggcggcc	tgcttgagga	tctcctgcac	gtcgcccgag	ttgtcctgat	aggcgatctc	7500
ggccatgaac	tggtccactt	cttcctcgcg	gaggtcaccg	tggcccgcctc	gctccacggt	7560
ggcgccagag	tcgttgagga	tgcggcgcac	gagttgagag	aaggcggtga	ggccgttctc	7620
gttccacacg	cggctgtaca	ccacgtttcc	gaaggagtcg	cgcgctcgca	tgaccacctg	7680
ggccacggtg	agttccacgt	ggcgggcgaa	gacggcgtag	tttctgaggc	gctggaagag	7740
gtagttgagc	gtggtggcga	tgtgctcgca	gacgaagaag	tacataatcc	agcgccgcag	7800
ggatcatctc	ttgatgtctc	cgatggcttc	gagacgctcc	atggcctcgt	agaagtcgac	7860
ggcgaagttg	aaaaattggg	agttgcgggc	ggccaccgtg	agttcttctt	gcaggaggcg	7920
gatgagatcg	gcgaccgtgt	cgcgcacctc	ctgttcgaaa	gcgcccgcag	gcgcctctgc	7980
ttcttcctcc	ggctcctcct	cttcaggggg	ctcggttcc	tccggcagct	ctgcgacggg	8040
gacggggcgg	cgacgtcgct	gtctgaccgg	caggcggtcc	acgaagcgct	cgatcatttc	8100
gccgcgccgg	cgacgcatgg	tctcggtgac	ggcgcgctcc	tttctcgag	gtcgcagttc	8160
gaagacgccg	ccgcgcagag	cgcccccg	caggggagggt	aagtgggttag	ggccgtcggg	8220
cagggacacg	gcgctgacga	tgcattttat	caattgctgc	gtaggcactc	cgtgcaggga	8280
tctgagaacg	tcgaggtcga	cgggatccga	gaacttctct	aggaaagcgt	ctatccaatc	8340
gcaatcgcaa	ggtaagctga	gaacggtggg	tcgctggggg	gcgttcgcgg	gcagttggga	8400
ggtgatgctg	ctgatgatgt	aattaaagta	ggcggtcttc	aggcggcgga	tggtggcgag	8460
gaggaccacg	tctttggggc	cggcctgttg	aatgcgcagg	cgctcggcc	tgccccaggc	8520
ctcgctctga	cagcgacgca	ggtctttgta	gaagtcttgc	atcagtctct	ccaccggaac	8580
ctctgcttct	cccctgtctg	ccatgcgagt	cgagccgaac	ccccgcagg	gctgcagcaa	8640
cgctaggctg	gccacgacct	tttcggccag	cacggcctgt	tgaatctgcg	tgagggtggc	8700
ctggaagtcg	tccaggtcca	cgaagcggtg	ataggcccc	gtgttgatgg	tgtaggtgca	8760
gttggccatg	acggaccagt	tgacgacttg	catgccgggt	tggtgatct	ccgtgtactt	8820
gaggcgcgag	taggccctgg	actcgaacac	gtagtcggtg	catgtgcgca	ccagatactg	8880
gtagccgacc	aggaagtga	gaggcggtc	tcggtacagg	ggccagccaa	cgggtggcggg	8940
ggcgccgggg	gacaggtcgt	ccagcatgag	gcggtggtag	tggtgatgt	agcgggagag	9000
ccaggtgatg	ccggccgagg	tggttgcggc	cctggtgaat	tcgcggacgc	ggttccagat	9060
gttgcgcgag	ggaccaaagc	gctccatggt	gggcacgctc	tgccccgtga	ggcgggcgca	9120
atcttgtagc	ctctagatgg	aaaaaagaca	ggcggtcat	cgactccttt	ccgtagcttg	9180
gggggtaaag	tcgcaagggt	gcggcggcgg	ggaaccccg	ttcgagaccg	gccggatccg	9240
ccgctcccga	tgcgcttggc	cccgcatcca	cgacgtccgc	gccgagacct	agccgcgacg	9300
ctccgcccc	atacgagggg	gagtcttttg	gtgttttttc	gtagatgcat	ccggtgctgc	9360
ggcagatgcg	accccagacg	cccactacca	ccgccgtggc	ggcagtaaac	ctgagcggag	9420
gcggtgacag	ggaggaggaa	gagctggctt	tagacctgga	agagggagag	gggctggccc	9480
ggctgggagc	gccatcccc	gagagacacc	ctagggttca	gctcgtgagg	gacgccaggc	9540

aggcttttgt gccgaagcag aacctgttta gggaccgcag cggtcaggag gcggaggaga	9600
tgcgcgattg caggtttcgg gcgggcagag agctcagggc gggcttcgat cgggagcggc	9660
tcctgagggc ggaggatttc gagcccgacg agcgttctgg ggtgagcccg gcccgcgctc	9720
acgtatcggc ggccaacctg gtgagcgcggt acgagcagac ggtgaacgag gagcgcaact	9780
tccaaaagag ctttaacaat cacgtgagga ccctgatcgc gagggaggag gtgaccatcg	9840
ggctgatgca tctgtgggac ttcgtggagg cctacgtgca gaacccggct agcaaacc	9900
tgacggccca gctgttcctg atcgtgcagc acagccgcga caacgagacg ttccgcgacg	9960
ccatgttgaa catcgcgag cccgagggtc gctggctctt ggatctgatt aacatcctgc	10020
agagcatcgt ggtgcaggag aggggcctga gtttagcggg caaggtggcg gccattaact	10080
attcgatgca gagcctgggg aagttctacg ctgcgaagat ctacaagagc ctttacgtgc	10140
ccatagacaa ggaggtgaag atagacagct tttacatgcg catggcgctg aaggtgctga	10200
cgctgagcga cgacctcggc gtgtaccgta acgacaagat ccacaaggcg gtgagcgcca	10260
gccgccggcg ggagctgagc gacagggagc tgatgcacag cctgcagagg gcgctggcg	10320
gcgccgggga cgaggagcgc gaggcttact tcgacatggg agccgatctg cagtggcgctc	10380
ccagcgcgcg cgcttggag gcggcggggt atcccgacga ggaggatcgg gacgatttgg	10440
aggaggcagg cgagtacgag gacgaagcct gaccgggcag gtgttggttt agatgcagcg	10500
gccggcggac gggaccaccg cggatcccg ctttttgga tccatgcaga gtcaaccttc	10560
gggcgtgacc gcctccgatg actgggcggc ggccatggac cgcacatcgg cgctgaccac	10620
ccgcaacccc gaggttttta ggcagcaacc ccaggccaac cgtttttcgg ccatcttgga	10680
agcggtggtg ccgtcgcgca ccaacccgac gcacgagaaa gtcctgacta tcgtgaacgc	10740
cctggtagac agcaaggcca tccgccgtga cgaggcggc ttgatttaca acgctctttt	10800
ggaacgcgtg gcgcgctaca acagcactaa cgtgcagacc aatctggacc gcctcaccac	10860
cgacgtgaag gaggcgctgg cgcagaagga gcggtttctg agggacagta atctgggctc	10920
tctggtagga ctgaacgcct tcctgagctc acagccggcc aacgtgcccc gcgggcagga	10980
ggattacgtg agcttcatca gcgctctgag actgctgggtg tccgaggtgc cccagagcga	11040
ggtgtaccag tctgggccgg attacttttt ccagacgtcc cgacagggct tgcaaacggt	11100
gaacctgact caggccttta aaaacttgca aggcattgtg ggggtcaagg ccccggtggg	11160
cgatcgcgcc actatctcca gtctgtgac cccaacact cgcctgctgc tgctcttgat	11220
cgcaccgttt accaacagta gcactatcag ccgtgactcg tacctgggtc atctcatcac	11280
tctgtaccgc gaggccatcg gccaggctca gatcgacgag catacgtatc aggagattac	11340
taacgtgagc cgtgccctgg gtcaggaaga taccggcagc ctggaagcca cgttgaactt	11400
tttgctaacc aaccggaggc aaaaaatacc ctcccagttc acgttaagcg ccgaggagga	11460
gaggattctg cgatacgtgc agcagtcctg gagcctgtac ttgatgcgcg agggcgccac	11520
cgcttccacg gcttttagaca tgacggctcg gaacatggaa ccgtcctttt actccgcca	11580

ccggccggttc	attaaccgtc	tgatggacta	cttccatcgc	gcggccgcca	tgaacgggga	11640
gtacttcacc	aatgccatcc	tgaatccgca	ttggatgccc	ccgtccgggt	tctacaccgg	11700
ggagtttgac	ctgcccgaag	ccgacgacgg	ctttctgtgg	gacgacgtgt	ccgatagcat	11760
tttcacgccc	gctaatacgcc	gattccagaa	gaaggagggc	ggagacgagc	tccccctctc	11820
cagcgtggaa	gcggcctcaa	ggggagagag	tccctttcca	agtctgtctt	ccgccagtag	11880
cggtcgggta	acgcgtccac	ggttgcccgg	ggagagcgac	tacctgaacg	accccttgct	11940
gcgaccggct	agaaagaaaa	attttcccaa	taacggggtg	gaaagcttgg	tggataaaat	12000
gaatcgttgg	aagacgtacg	cccaggagca	gcgggagtgg	gaggacagtc	agccgcggcc	12060
gctggtaccg	ccgcattggc	gtcgccagag	agaagacccg	gacgactccg	cagacgatag	12120
tagcgtgttg	gacctgggag	ggagcggagc	caaccccttt	gctcacttgc	aaccaagggg	12180
gcgctcgagt	cgctgtatt	aataaaaaag	acgcggaaac	ttaccagagc	catggccaca	12240
gcgtgtgtgc	tttcttctc	tctttcttcc	tcggcgcggc	agaatgagaa	gagcggtgag	12300
agtcacgccc	gcggcgtatg	agggcccgcc	cccttcttac	gaaagcgtga	tgggatcagc	12360
gaacgtgccg	gccacgtgg	aggcgcctta	cgcttctccc	agatacctgg	gacctacgga	12420
gggcagaaac	agcatccgtt	actccgagct	ggcgcccctg	tacgatacca	ccaagggtga	12480
cctggtggac	aacaagtcgg	cggacatcgc	ctccctgaat	taccaaaccg	atcacagtaa	12540
ctttctgact	accgtggtgc	agaacaatga	cttcaccccg	acggaggcg	gcacgcagac	12600
cattaacttt	gacgagcggt	cccgtgggg	cggtcagctg	aaaaccatcc	tgcacaccaa	12660
catgccaac	atcaacgagt	tcatgtccac	caacaagttc	agggctaagc	tgatggtaga	12720
aaaaagtaat	gcggaaactc	ggcagccccg	atacagtggt	ttcgagttta	ccattccaga	12780
gggcaactat	tccgaaacta	tgactatcga	tctcatgaat	aacgcgatcg	tggacaatta	12840
cctgcaagtg	gggagacaga	acggggtgct	ggaaagcgat	atcggcgtga	aattcgatac	12900
cagaaacttc	cgactggggg	gggatcccgt	gaccaagctg	gtgatgccag	gcgtgtacac	12960
caacgagggt	tttcacccgg	acatcggtg	gctgccgggg	tgcggtgtgg	acttcaactca	13020
gagccgtttg	agtaacctgt	taggaattag	aaagcgccgc	cccttccaag	agggttttca	13080
aatcatgtat	gaggacctgg	agggaggtta	tatacccgcc	ttactggacg	tgtcgaagta	13140
cgaagctagc	atacaacgcg	ccaaagcgga	gggtagagag	attcggggag	acacctttgc	13200
ggtagctccc	caggacctgg	aaatagtgcc	tttaactaaa	gacagcaaag	acagaagcta	13260
caatattata	aacaacacga	cggacaccct	gtatcgagac	tggtttctgg	cttacaacta	13320
cggagacccc	gagaaaggag	tgagatcatg	gaccatactc	accaccacgg	acgtgacctg	13380
tggctcgag	caagtgtact	ggtccctgcc	ggatatgatg	caagacccgg	tcaccttccg	13440
cccctccacc	caagtcagca	acttcccggg	gggtgggcacc	gagctgctgc	ccgtccatgc	13500
caagagcttc	tacaacgagc	aggccgtcta	ctcgcaactt	attcgccagt	ccaccgcgct	13560
taccacagtg	ttcaatcgct	ttcccagaaa	ccagattctg	gtgcgccctc	ccgtccttac	13620
cattaccacc	gtcagtgaag	acgttcccgc	cctcacagat	cacggaaccc	tgccgtgcg	13680

cagcagtatc agtggagttc agcgcgtgac catcaccgac gccagacgtc gaacctgccc 13740
ctacgtttac aaagcgcttg gcgtggtggc tcctaaagtt ctttctagtc gcaccttcta 13800
aaaacatgtc catcctcatc tctcccgata acaacaccgg ctggggactg ggctccggca 13860
agatgtacgg cggagccaaa aggcgctcca gtcagcacc agttcgagtt cggggccact 13920
tccgcgctcc ttggggagct tacaagcgag gactctcggg tcgaacggct gtagacgata 13980
ccatagatgc cgtgattgcc gacgcccgcc ggtacaaccc cggaccggtc gctagcgccg 14040
cctccaccgt ggattccgtg atcgacagcg tggtagccgg cgctcgggccc tatgctcgcc 14100
gcaagaggcg gctgcatcgg agacgtcgcc ccaccgccgc catgctggca gccagggccg 14160
tgctgaggcg ggcccggagg gcaggcagaa gggctatgcg ccgcgctgcc gccaacgccg 14220
ccgccgggag ggcccgccga caggctgccc gccaggctgc cgctgccatc gctagcatgg 14280
ccagaccag gagagggaac gtgtactggg tgcgtgattc tgtgacggga gtccgagtgc 14340
cgggtgcgag ccgacctccc cgaagttaga agatccaagc tgcgaagacg gcggtactga 14400
gtctccctgt tgttatcagc ccaacatgag caagcgcaag tttaaagaag aactgctgca 14460
gacgctggtg cctgagatct atggccctcc ggacgtgaag ccagacatta agccccgcga 14520
tatcaagcgt gttaaaaagc gggaaaagaa agagggaactc gcggtggtag acgatggcgg 14580
agtggaattt attaggagtt tcgccccgcg acgcagggtt caatggaaag ggcggcggg 14640
acaacgcgtt ttgaggccgg gcaccgcggt agtttttacc ccgggagagc ggtcggccgt 14700
taggggtttc aaaaggcagt acgacgaggt gtacggcgac gaggacatat tggaacaggc 14760
ggctcaacag atcggagaat ttgcctacgg aaagcgttcg cgtcgcgaag acctggccat 14820
cgccttagac agcggcaacc ccacgcccag cctcaaacc gtgacgctgc agcagggtgct 14880
tcccgtgagc gccagcacgg acagcaagag ggggattaag agagaaatgg aagatctgca 14940
tcccaccatc caactcatgg tccctaaacg gcagaggctg gaagagggtc tggagaagat 15000
gaaagtggac cccagcatag agccggatgt aaaagtcaga cctattaagg aagtggcccc 15060
cggctcttggg gtgcaaacgg tggacattca aatccccgtc accaccgctt caaccgccgt 15120
ggaagctatg gaaacgcaaa cggagacccc tgccgcgac ggtaccaggg aagtggcggt 15180
gcaaacggag ctttggtagc aatacgagc ccctcggcgt cagaggcggt ccgctcggtta 15240
cggccccgcc aacgccatca tgccagaata tgcgctgcat ccgtctattc tgccccactc 15300
cggataccgg ggtgtgacgt atcgcccgtc tggaaaccgc cgccgaaccc gtcgccgccg 15360
ccgctcccgt cgcgctctgg cccccgtgtc ggtgcggcgt gtgaccgccg ggggaaagac 15420
agtcgtcatt cccaaccgc gttaccaccc tagcatcctt taataactct gccgttttgc 15480
agatggctct gacttgccgc gtgcgccctc ccgttcgca ctatcgagga agatctcgtc 15540
gtaggagagg catgacgggc agtggctgcc ggcgggcttt gcgcaggcgc atgaaaggcg 15600
gaattttacc cgccctgata ccataattg ccgccgccat cggtgccata cccggcggtg 15660
cttcagtggc gttgcaagca gctcgtaata aataaaciaa ggctttttgca cttatgacct 15720

ggtcctgact attttatgca gaaagagcat ggaagacatc aattttacgt cgctggctcc 15780
gcggcacggc tcgcggccgc tcatgggcac ctggaacgac atcggcacca gtcagctcaa 15840
cgggggcgct ttcaattggg ggagcctttg gagcggcatt aaaaactttg gctccacgat 15900
taaatcctac ggcagcaaag cctggaacag tagtgctggt cagatgctcc gagataaact 15960
gaaggacacc aacttccaag aaaaagtggg caatgggggtg gtgaccggca tccacggcgc 16020
ggtagatctc gccaaccaag cgggtgcagaa agagattgac aggcgttttg aaagctcgcg 16080
gggtgccgccg cagagagggg atgaggtgga ggctgaggaa gtagaagtag aggaaaagct 16140
gccccgcgtg gagaaagttc ccggtgcgcc tccgagaccg cagaagcggc ccaggccaga 16200
actagaagag actctggtga cggagagcaa ggagcctccc tcgtacgagc aagccttgaa 16260
agagggcgcc tctccaccct cctacccgat gactaagccg atcgcaccca tggctcgacc 16320
gggtgtacggc aaggattaca agcccgtcac gctagagctg cccccaccgc cccccacgcg 16380
cccgaccgtc cccccctgc cgactccgtc ggcgcccgcg gcgggaccgc tgtccgcacc 16440
atccgctgtg cctctgccag ccgcccgtcc agtggccgtg gccactgcca gaaaccccag 16500
aggccagaga ggagccaact ggcaaagcac gctgaacagc atcgtgggcc tgggagtga 16560
aagcctgaaa cgccgccgtt gctattatta aaaaagtgtg gctaaaaagt ctcccgttgt 16620
atacgctcc tatgttaccg ccagagacga gtgactgtcg ccgcgagcgc cgctttcaag 16680
atggccaccc catcgatgat gccgcagtgg tcttacatgc acatcgccgg ccaggacgcc 16740
tcggagtacc tgagtcccg cctcgtgcag tttgcccgcg ccaccgacac ctacttcagc 16800
ttgggaaaca agtttagaaa ccccaccgtg gccccaccc acgatgtgac cacggaccgc 16860
tcgcagaggc tgaccctgcg ctttgtgccc gtagaccggg aggacaccgc gtactcttac 16920
aaagtgcgct acacgttggc cgtaggggac aaccgagtgc tggacatggc cagcacctac 16980
tttgacatcc ggggggtgct ggatcggggc cccagcttca agccctattc cggcaccgct 17040
tacaactccc tggcccccaa gggagctccc aaccctcgg aatggacgga cacttccgac 17100
aacaactta aagcatatgc tcaggctccc taccagagtc aaggacttac aaaggatggg 17160
attcaggttg ggctagtgtg gacagagtca ggacaaacac cccaatatgc aaacaaagt 17220
taccaacccg agccacaaat tggggaaaac caatggaatt tagaacaaga agataaagcg 17280
gcgggaagag tcctaaagaa agatacccct atgtttccct gctatgggtc atatgccagg 17340
cccacaaacg aacaaggagg gcaggcaaaa aaccaagaag tagatttaca gttttttgcc 17400
actccgggcg acaccagaa cacggctaaa gtggtacttt atgctgaaaa tgtcaacctg 17460
gaaactccag atactcactt agtgtttaaa cccgatgacg acagcaccag ttcaaaactt 17520
cttcttgggc agcaggctgc acctaacaga cccaactaca taggttttag agataatttt 17580
attggtttta tgtactacaa tagcactgga aacatgggcg tgctggccgg acaggcttct 17640
caattgaatg ccgtagtcga cttgcaggac agaaacaccg agttgtccta ccagctgatg 17700
ctggacgcac tgggggatcg cagccgatat ttttcaatgt ggaatcaggc agtagacagc 17760
tatgaccag acgttagaat tatagaaaac cacggagtgg aagacgaact gccaaactat 17820

tgttttcctc	tgggaggaat	ggtggtgact	gacaattaca	actctgtgac	gcctcaaaat	17880
ggaggcagtg	gaaatacatg	gcaggcagac	aatactacat	ttagtcaaag	aggagcgcag	17940
attggctccg	gaaacatggt	tgccctggaa	attaacctac	aggccaacct	ctggcgcggc	18000
ttcttgat	ccaatattgg	gttgatctt	ccagactctc	tgaaaatcac	ccccgacaac	18060
atcacgctgc	cagaaaacaa	aaacacttat	cagtacatga	acggtcgcgt	aacgccaccc	18120
gggctcatag	acacctatgt	aaacgtgggc	gcgcgctggt	ccccgatgt	catggacagc	18180
attaaccctt	tcaaccacca	ccgtaacgcg	ggcttgcgct	accgctccat	gctcttgggc	18240
aacggccggt	atgtgccttt	tcacattcag	gtgccccaaa	aattctttgc	cattaaaaac	18300
ctgctgcttc	tccccgggtc	ctatacctat	gagtggaaact	tccgcaagga	tgtcaacatg	18360
atcctgcaga	gctcgctggg	taatgacctg	cgagtggacg	gggccagcat	acgctttgac	18420
agcattaacc	tgtatgccaa	cttttttccc	atggcccaca	acacggcctc	taccctggaa	18480
gccatgctgc	gcaacgacac	caatgaccag	tccttcaacg	actacctgtg	cgcggtaac	18540
atgctgtacc	ccatccccgc	caacgccacc	agcgtgccca	tttctattcc	ttctcggaac	18600
tgggctgcct	tcaggggctg	gagttttact	cgctcaaaa	ccaaggagac	tccctcgctg	18660
ggctccggtt	ttgaccctta	ctttgtttac	tccggctcca	ttccctacct	agatggcacc	18720
ttttacctca	accacacttt	caaaaagggtg	tctattatgt	ttgactcctc	ggttagctgg	18780
cccggcaacg	accgcctgct	aacgccaac	gagttcgaaa	ttaagcggtc	cgtggacggt	18840
gaaggggtaca	acgtggccca	gagcaacatg	accaaggact	ggtttctaata	tcaaagtctc	18900
agtcactata	atatagggtta	ccagggcttc	tatgtgcccg	agaactacaa	ggaccgcatg	18960
tactccttct	tccgcaactt	ccaaccaatg	agccggcagg	tggtagatac	cgtgacttat	19020
acagactaca	aagatgtcaa	gctcccctac	caacacaaca	actcaggggt	cgtgggctac	19080
atgggacca	ccatgcgaga	gggacaggcc	taccggcca	actatcccta	ccccctgatc	19140
ggagagactg	ccgtaccag	cctcacgcag	aaaaagttcc	tctgcgaccg	ggtgatgtgg	19200
aggataccct	tctctagcaa	ctttatgtcg	atgggctccc	tcaccgacct	ggggcagaac	19260
atgctgtacg	ccaactccgc	tcacgccttg	gacatgactt	ttgaggtgga	tcccatggat	19320
gagcccacgc	ttctctatgt	tctgtttgaa	gtcttcgacg	tggtgcgcat	ccaccagccg	19380
caccgcggcg	tcatcgaggc	cgtctacctg	cgcacacctt	tctctgccgg	taacgccacc	19440
acctaaagaa	gctgatgggt	tccagcgaac	aggagttgca	ggccattggt	cgcgacctgg	19500
gctgcggggc	ctgctttttg	ggcaccttcg	acaagcggtt	tcccggaattc	atgtcccccc	19560
acaagccggc	ctgcgccatc	gttaacacgg	ccggacggga	gacagggggg	gtgactggc	19620
tcgccttcgc	ctggaacccg	cgcaaccgca	cctgctacct	gttcgaccct	tttggtttct	19680
ccgacgaaag	gctgaagcag	atctaccaat	tcgagtacga	ggggctcctc	aagcgacgcg	19740
ctctggcctc	cacgcccagc	cactgcgtca	ccctggaaaa	gtccaccag	acggtccagg	19800
ggccccctc	ggccgcctgc	gggcttttct	gttgcatggt	tttgacgcc	ttcgtgcact	19860

ggcctcacac	ccccatggag	cgcaacccca	ccatggatct	gctcaccgga	gtgcccaca	19920
gcatgcttca	cagtccccag	gtcgccccca	ccctgcgtcg	caatcaggac	cacctgtatc	19980
gctttctggg	gaaacactct	gcctatttcc	gccgccaccg	gcagcgcata	gaacaggcca	20040
cggccttcga	aagcatgagc	caaagagtgt	aatcaataaa	aaccgttttt	atttgacatg	20100
atacgcgctt	ctggcgtttt	tattaaaaat	cgaagggttc	gagggagggg	tcctcgtgcc	20160
cgctggggag	ggacacgttg	cggtactgga	atcgggcgct	ccaacgaaac	tcggggatca	20220
ccagccgcgg	cagggccacg	tcttccatgt	tctgcttcca	aaactgtcgc	accagctgca	20280
gggctcccat	cacgtcgggc	gctgagatct	tgaagtcgca	gttagggccg	gagccccgcg	20340
ggctgttgcg	gaacacgggg	ttggcacact	ggaacaccaa	cacgctgggg	ttgtggatac	20400
tagccagggc	cgtcgggtcg	gtcacctccg	atgcatccag	atcctcggca	ttgctcaggg	20460
cgaacggggg	cagcttgac	atctgccgcc	cgatctgggg	taccaggtcg	cgcttggtga	20520
ggcagtcgca	gcgcagaggg	atgaggatgc	gacgctgcc	gcgttgcatg	atggggtaac	20580
tcgccgccag	gaactcctct	atctgacgga	aggccatctg	ggccttgacg	ccctcgggtga	20640
aaaatagccc	acaggacttg	ctggaaaaca	cgttattgcc	acagttgatg	tcttccgcgc	20700
agcagcgcgc	atcttcgttc	ttcagctgaa	ccacgttgcg	accccagcgg	ttctgaacca	20760
ccttggtttt	cgtgggatgc	tccttcagcg	cccgtgtcc	gttctcgtg	gtcacatcca	20820
tttccaccac	gtgctccttg	cagaccatct	ccactccgtg	gaaacagaac	agaatgccct	20880
cctgttgggg	attgcatg	tcccacacgg	cgcacccggt	ggactcccag	ctcttggtt	20940
tcacccccgc	gtaggcttcc	atgtaagcca	ttagaaatct	gcccatacag	tcagtgaagg	21000
tcttctgggt	ggtgaagggt	agcggcaggg	cgcggtgttc	ctcgttcaac	caagtttgac	21060
agatcttgcg	gtacacggct	ccctggtcgg	gcagaaactt	aaaagtcgtt	ctgctctcgt	21120
tgtccacgtg	gaacttctcc	atcaacatcg	tcatgacttc	catgcccttc	tcccaggcag	21180
tcaccagcgg	cgcgctctcg	gggttcttca	ccaacacggc	ggtggagggg	ccctcgccgg	21240
ccccgacgtc	cttcatggac	attttttgaa	actccacggt	gccgtccgcg	cggcgctactc	21300
tgcgcatcgg	agggtagctg	aagcccacct	ccatgacggt	gctttcgccc	tcgctgtcgg	21360
agacgatctc	cggggagggc	ggcggaacgg	gggcagactt	gcgagccttc	ttcttgggag	21420
ggagcggagg	cacctcctgc	tcgcgctcgg	gactcatctc	ccgcaagtag	ggggtgatgg	21480
agcttcctgg	ttggttctga	cggttggcca	ttgtatccta	ggcagaaaga	catggagctt	21540
atgcgcgagg	aaactttaac	cgccccgtcc	cccgtcagcg	acgaagaggt	catcgtcgaa	21600
caggacccgg	gctacgttac	gccgcccag	gatctggagg	ggcccttaga	cgaccggcgc	21660
gacgctagt	agcggcagga	aaatgagaaa	gaggaggagg	agggtgcta	cctcctggaa	21720
ggcgacgttt	tgctaaagca	tttcgccagg	cagagcacca	tactcaagga	ggccttgcaa	21780
gaccgctccg	aggtgccctt	ggacgtcgcc	gcgctctccc	aggcctacga	ggcgaacctt	21840
ttctcgcccc	gagtgcctcc	gaagagacag	cccaacggca	cctgcgagcc	caaccgcgca	21900
ctcaacttct	accccggtgt	cgccgtgccc	gaggcgctgg	ccacctacca	catctttttc	21960

aaaaaccagc gcattcccct ttcctgccgg gccaaaccgca ccgcggccga taggaagcta 22020
acactcagaa acggagtcag catacctgat atcacgtcac tggaggaagt gcctaagatc 22080
ttcgaggggtc tgggtcgaga tgagaagcgg gcggcgaaacg ctctgcagaa agaacagaaa 22140
gagagtcaga acgtgctggt ggagctggag ggggacaacg cgcgtctgac cgtcctcaaa 22200
cgttgcatag aagtttccca cttcgcctac ccggccctca acctgccgcc caaagttatg 22260
aaatcgggtca tggaccagct actcatcaag agagctgagc ccctgaatcc cgaccacctt 22320
gaggcggaaa actcagagga cggaaagccc gtcgtcagcg acgaggagct cgagcgggtg 22380
ctggaaacca gggacccccca gcagttgcaa gagaggcgca agatgatgat ggcggccgtg 22440
ctggtcacgg tggagctaga atgcctgcaa cggtttttca gcgacgtgga gacgctacgc 22500
aaaatcgggg agtccctgca ctacaccttc cgccagggtt acgttcgcca ggcctgcaaa 22560
atctccaacg tagagctcag caacctgggtt tcctacatgg gcatcctcca cgagaaccgg 22620
ctggggcaga gcgtgctgca ctgcacctg caaggcgagg cggaaggga ctacgtccga 22680
gactgcgctt acctcttcct caccctcacc tggcagaccg ccatgggctg gtggcagcag 22740
tgcttgggaag agagaaacct caaagagctg gacaaactcc tctgccgcca gcggcgggccc 22800
ctctggaccg gcttcagcga gcgcacgggtc gcctgcgccc tggcagacat cattttccca 22860
gaacgcctga tgaaaacctt gcagaacggc ctgccggatt tcatcagtca gagcatcttg 22920
caaaacttcc gctccttcgt cctggagcgc tccgggatct tgcccgccat gagctgcgcg 22980
ctgccttctg actttgtccc ctttctctac cgcgagtgcc ctccccact gtggagccac 23040
tgctacctct tccaactggc caactttctg gcctaccact ccgacctcat ggaagacgtg 23100
agcggagagg ggctgctcga gtgccactgc cgctgcaacc tctgcacccc ccacagatcg 23160
ctggcctgca acaccgagct gctcagcgaa acccagggtca taggtacctt cgagatccag 23220
gggccccagc agcaagaggg tgcttccggc ttgaagctca ctccggcgct gtggacctcg 23280
gcttacttac gcaaatttgt agccgaggac taccacgccc acaaaattca gttttacgaa 23340
gaccaatctc gaccaccgaa agccccctc acggcctgcg tcatcaccca gagcaaaatc 23400
ctggcccaat tgcaatccat caaccaagcg cgccgagatt tccttttgaa aaagggctcg 23460
ggggtgtacc tggacccccca gaccggcgag gaactcaacc cgtccacact ttccgtcgaa 23520
gcagcccccc cgagacatgc caccgaagg aaccgccaag cagctgatcg ctcggcagag 23580
agcgaagaag caagagctgc tccagcagca ggtggaggac gaggaagagc tgtgggacag 23640
ccaggcagag gaggtgtcag aggacgagga ggagatggaa agctgggaca gcctagacga 23700
ggaggacgag ctttcagagg aagaggcgac cgaagaaaaa ccacctgcat ccagcgcgcc 23760
ttctctgagc cgacagccga agccccggcc ccgcagccc ccggccggct cactcaaagc 23820
cagccgtagg tgggacgcca ccggatctcc agcggcagcg gcaacggcag cgggtaaggc 23880
caaacgcgag cggcgggggt attgctcctg gcggaccac aaaagcagta tcgtgaactg 23940
cttgcaacac tgcgggggaa acatctcctt tgcccagcgc tacctcctct tccatcacgg 24000

tgtggccttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24060
gctcggagaa aaaagctaag gcctcctctg ccgcgaggaa aaactccgcc gccgctgccg 24120
ccaaggatcc gccggccacc gaggagctga gaaagcgcac ctttcccact ctgtatgcta 24180
tctttcagca aagccgcggg cagcaccctc agcgcgaact gaaaataaaa aaccgctcct 24240
tccgctcact caccgcagc tgtctgtacc acaagagaga agaccagctg cagcgcaccc 24300
tggacgacgc cgaagcactg ttcagcaaact actgctcagc gtctcttaaa gactaaaaga 24360
cccgcgcttt ttccccctcg ggcgccaaaa ccacgtcat cgccagcatg agcaaggaga 24420
ttcccacccc ttacatgtgg agctatcagc ccagatggg cctggccgcg ggggcccgcc 24480
aggactactc cagcaaaatg aactggctca gcgccggccc ccacatgatc tcacgagtta 24540
acggcatccg agcccaccga aaccagatcc tcttagaaca ggcggaatc accgccacac 24600
cccggcgcca actcaaccg cccagttggc ccgccgcca ggtgtatcag gaaactcccc 24660
gcccgaccac agtcctcctg ccacgcgacg cggaggccga agtcctcatg actaactctg 24720
gggtacaatt agcggggcggg tccaggtacg ccaggtacag aggtcgggccc gctccttact 24780
ctcccgggag tataaagagg gtgatcattc gaggccgagg tatccagctc aacgacgagg 24840
cggtgagctc ctcaaccggt ctacagctg acggagtctt ccagctcgga ggagcgggccc 24900
gctcttcctt caccactcgc caggcctacc tgaccctgca gagctcttcc tcgcagccgc 24960
gctccggggg aatcggcact ctccagttcg tggaagagtt cgtcccctcc gtctacttca 25020
acccgttttc cggctcacct ggacgctacc cggacgcctt cattcccaac tttgacgcag 25080
tgagtgaatc cgtggacggc tacgactgat gacagatggg gcggccgtga gagctcggct 25140
gcgacatctg catcactgcc gccagcctcg ctgctacgct cgggaggcga tcgtgttcag 25200
ctactttgag ctgccggacg agcaccctca gggaccggct cacgggttga aactcgagat 25260
tgagaacgcg cttgagtctc acctcatcga cgccttcacc gcccggcctc tcctggtaga 25320
aaccgaacgc gggatcacta ccatcacct gttctgcatc tgccccacgc ccggattaca 25380
tgaagatctg tgttgtcatc tttgcgctca gtttaataaa aactgaactt tttgccgtac 25440
cttcaacgcc acgcgttggt tctccttggt aaaaaacccc aggagtctt aacttacaca 25500
tagcaaaacc cttgtatttt accatagaaa aacaactagc cttttcaatt ggaaaagggg 25560
taacaatttc tgctacagga cagttgaaa gcacagcaag cgtacaggac agcgtacac 25620
caccctacg tggattttcc cttttaagc tgacagacaa cggtttaaca ttaagctatt 25680
cagatcccct gcgtgtggt ggtgaccaac ttacgtttaa ttttacttct ccactacggt 25740
acgaaaatgg cagtcttaca ttcaactaca cttctcccat gacactaata aacaacagtc 25800
ttgctattaa cgtcaatacc tccaaaggcc tcagtagtga caacggcaca ctgcgtgtaa 25860
atgttactcc agattttaga tttaacagct ctggtgcctt aacttttggc atacaaagtc 25920
tatggacttt tccaacaaa actcctaact gtaccgtgtt taccgaaagt gactccctgc 25980
tgagtctttg cttgactaaa tgcggagctc acgtacttgg aagcgtgagt ttaagcggag 26040
tggcaggaac catgctaaaa atgaccaca cttctgttac cgttcagttt tcgtttgatg 26100

acagtggtaa actaatattc tctccacttg cgaacaacac ttgggggtggt cgacaaagcg 26160
 agagtccgtt gccaaccca tccttcaacg ctctcacgtt tatgccaaac agtaccattt 26220
 attctagagg agcaagtaac gaacctcaaa acaattatta tgtccagacg tatcttagag 26280
 gcaacgtgcg aaagccaatt ctactaactg ttacctacaa ctgagttaat tcaggatatt 26340
 ccttaacttt taaatgggat gctgtcgcca atgaaaaatt tgccactcct acatcttcgt 26400
 tttgctatgt tgcagagcaa taaaacctg ttacccacc gtctcgtttt tttcagatga 26460
 aacgagcgag agttgatgaa gacttcaacc cagtgtacc ttatgacccc ccatacgctc 26520
 ccgtcatgcc cttcattact ccgcctttta cctcctcgga tgggttgag gaaaaaccac 26580
 ttggagtgtt aagtttaaac tacagggatc ccattactac acaaaatggg tctctcacgt 26640
 taaaactagg aaacggcctc actctaaaca accagggaca gttaacatca actgctggcg 26700
 aagtggagcc tccgctcact aatgctaaca acaaaactgc actagcctat agcgaaccat 26760
 tagcagtaaa aagcaaccgc ctaactctat cacacaccgc tccccttgct atcgctaata 26820
 attcttttagc gttgcaagtt tcagagccta tttttgtaaa tgacgatgac aagctagccc 26880
 tgcagacagc cgccccctt gtaaccaacg ctggcaccct tcgcttacag agcgctgccc 26940
 ctttaggatt ggttgaaaat actcttaaac tgctgttttc taaacccttg tatttgcaaa 27000
 atgattttct tgcattagcc attgaacgcc ccctggctgt agcagccgca ggtactctga 27060
 ccctacaact tactcctcca ttaaagacta acgatgacgg gctaacta tccacagtcg 27120
 agccattaac tgtaaaaaac ggaaacctag gcttgcaa atcgcgccct ttagttgttc 27180
 aaaacaacgg ctttcgctt gctattacc ccccgctgcg tttgtttaac agcgaccccg 27240
 ttcttggttt gggcttact tttcccctag ctgtcaciaa caacctctc tccttaaaaca 27300
 tgggagacgg agttaaactt acctataata aactaacagc caatttgggt agggatttac 27360
 aatttgaaaa cgggtgcgatt gccgtaacgc ttactgccga attacctttg caataacta 27420
 acaaaactca actgaatatt ggagctggcc ttcgttacaa tggagccagc agaaaactag 27480
 atgtaaacat taacaaaaat aaaggcttaa cttgggacaa cgatgcagtt attcccaaac 27540
 taggatcggg cttacaattt gaccctaata gcaacatcgc tgttatccct gaaaccgtga 27600
 agccgcaaac gttatggacg actgcagatc cctcgcctaa ctgctcagtg taccaggact 27660
 tggatgccag gctgtggctc gctcttgcta aaagtggcga catggtgcat ggaagcattg 27720
 ccctaaaagc ctaaaaaggg acgttgctaa atcctacagc cagctacatt tccattgtga 27780
 tatattttta cagcaacgga gtcaggcgta ccaactatcc aacgtttgac aacgaaggca 27840
 ccttagctaa cagcgccact tggggatacc gacaggggca atctgctaac actaatgtga 27900
 ccaatgccac tgaatttatg ccagctcaa gcaggtacc cgtgaataaa ggagacaaca 27960
 ttcaaaatca atctttttca tacacctgta ttaaaggaga ttttgctatg cctgtcccgt 28020
 tccgtgtaac atataatcac gccctggaag ggtattccct taagttcacc tggcgcggtg 28080
 tagccaatca ggcctttgat attccttgct gttcattttc atacatcaca gaataaaaaa 28140

ccacttttttc attttaatttt cttttttatttt tacacgaaca gtgagacttc ctccaccctt 28200
ccatttgaca gcatacacca gcctctcccc cttcatagca gtaaactgtt gtgaatcagt 28260
ccggtattttg ggagttaaaa tccaaacagt ctcttttggtg atgaaacgtc gatcagtaat 28320
ggacacaaat ccctgggaca ggtttttccaa cgtttcggtg aaaaactgca caccgcccta 28380
caaaacaaac aggttcaggc tctccacggg ttatctcccc gatcaaactc agacagggtgta 28440
aagggtgcggt ggtgtttccac taaaccacgc aggtggcgct gtctgaacct ctcggtgcga 28500
ctcctgtgag gctggtaaga agttagattg tccagtagcc tcacagcatg tatcatcagt 28560
ctacgagtgc gtctggcgca gcagcgcac tgaatctcac tgagattccg gcaagaatcg 28620
cacaccatca caatcagggtt gttcatgatc ccatagctga acacgctcca gccaaagctc 28680
attcgctcca acagcgccac cgcggtgtccg tccaacctta ctttaacata aatcagggtgt 28740
ctgccgcgta caaacatgct acccacatc agaacttccc ggggcaggcc cctgttcacc 28800
acctgtctgt accagggaac cctcacattt atcaggggagc catagatggc cattttaaac 28860
caattagcta ataccgcccc accagctcta cactgaagag aaccgggaga gttacaatga 28920
cagtgaataa tccatctctc ataaccctg atggtctgat gaaaatctag atctaacgtg 28980
gcacaacaaa tacacacttt catatacatt ttcataacat gttttttcca ggccgttaaa 29040
atacaatccc aatacacggg ccactcctgc agtacaataa agctaataca agatggtata 29100
ctcctcacct cactgacact gtgcatgttc atattttcac attctaagta ccgagagttc 29160
tcctctacag cagcactgct gcggtcctca caagggtgta gctgggtgatg attgtagggg 29220
gccagtctgc agcgataccg tctgtcgcgt tgcacgtag accaggaacc gacgcacctc 29280
ctcgtacttg tggtagcaga accacgtccg ctgccagcac gtctccacgt aacgccggtc 29340
cctgcgtcgc tcacgtccc tcctcaatgc aaagtgaac cactcttgta atccacacag 29400
atccctctcg gcctccgggg tgatgcacac ctcaaacta cagatgtctc ggtacagttc 29460
caaacacgta gtgagggcga gttccaacca agacagacag cctgatctat cccgacacac 29520
tgagggtgga ggaagacacg gaagaggcat gttattccaa gcgattcacc aacgggtcga 29580
aatgaagatc ccgaagatga caacgggtcgc ctccggagcc ctgatggaat ttaacagcca 29640
gatcaaactg tatgcgattc tccaagctat cgatcgccgc ttccaaaaga gcctggaccc 29700
gcacttccac aaacaccagc aaagcaaaag cactattatc aaactcttca atcatcaagc 29760
tgcaggactg tacaatgcct aagtaatttt cgtttctcca ctgcgcaatg atgtcgcggc 29820
agatagtctg aaggttcac cctgacggg taaaagctc cgaaagggcg ccctctacag 29880
ccatgcgtag acacaccatc atgactgcaa gatatcgggc tcctgagaca cctgcagcag 29940
atttaacaga tcaaggctcag gttgctctcc gcgatcacga atctccatcc gcaaggctcat 30000
ttgcaaaaaa ttaaataaat ctatgccgac tagatctgtc aactccgcat taggaaccaa 30060
atcagggtgtg gctacgcagc acaaaagttc cagggatggt gccaaactca ctagaaccgc 30120
tcccagagtaa caaaactgat gaatgggagt aacacagtgt aaaatgtgca accaaaaatc 30180
actaagggtgc tccttttaaaa agtccagtac ttctatattc agtccgtgca agtactgaag 30240

```

caactgtgcg ggaatatgca caacaaaaaa aatagggcgg ctgagatata tgttgacctt 30300
aaataaaaaag aatcattaaa ctaaagaagc ttggcgaacg gtgggataaa tgacacgctc 30360
cagcagcaga caggcaaccg gctgtccccg ggaaccgcgg taaaattcat ccgaatgatt 30420
aaaaagaaca acagaaactt cccaccatgt actcggttgg atctcctgag cacacagcaa 30480
tccccccctc acattcatgt ccgccacaga aaaaaaacgt cccagatacc cagcggggat 30540
atccaacgac agctgcaaag acagcaaaac aatccctctg ggagcgatca caaaatcctc 30600
cgggtgaaaaa agcacatata tattagaata accctgttgc tggggcaaaa agggccggcg 30660
tcccagcaaa tgcacataaa tatgttcatc agccattgcc ccgtcttacc gcgtaatcag 30720
ccacgaaaaa atcgagctaa aattcaccca acagcctata gctatatata cactccgccc 30780
aatgacgcta ataccgcacc acccacgacc aaagttcacc cacaccaca aaaccgcgca 30840
aatccagcg ccgtcagcac ttccgcaatt tcagtctcac aacgtcactt ccgcgcgcct 30900
tttcacattc ccacacacac ccgcgccctt cggcccgccc tcgcgccacc ccgcgtcacc 30960
gcacgtcacc ccggccccgc ctgcgtcctc cccgctcatt atcatattgg cacgtttcca 31020
gaataaggta tattattgat gatg 31044

```

<210> 6

<211> 34115

<212> DNA

<213> simian adenovirus SV-39

<400> 6

```

catcatcaat ataacaccgc aagatggcga ccgagttaac atgcaaataa ggtgggcgga 60
gttacgcgac ctttgtcttg ggaacgcgga agtgggcgcg gcgggtttcg gggaggagcg 120
cggggcgggg cgggcgtgtc gcgcggcggt gacgcgccgg ggaccggaa attgagtagt 180
ttttattcat tttgcaagtt tttctgtaca ttttggcgcg aaaactgaaa cgagggaagt 240
aaaagtgaaa aatgccgagg tagtcaccgg gtggagatct gacctttgcc gtgtggagtt 300
taccgcgtga cgtgtgggtt tcggtctcta ttttttact gtggttttcc gggtagcggtc 360
aaaggtcccc attttatgac tccacgtcag ctgatcgcta gggatattta tgccctcag 420
accgtcaaga ggccactctt gagtgcgggc gagaagagtt ttctcctccg cgttccgcca 480
actgtgaaaa aatgaggaac ttcttgctat ctccggggct gccagcgacc gtagccgccc 540
agctgttggg ggacattggt accggagctc tgggagacga tcctcagggtg atttctcact 600
tttgtgaaga ttttagtctt catgatctct atgatattga tccgggtggt gaggggcaag 660
aggatgaatg gctggagtct gtggatgggt tttttccgga cgctatgctg ctagaggctg 720
atttgccacc acctcacaac tctcacactg agcccgagtc agctgctatt cctgaattgt 780
catcaggtga acttgacttg gcttgttacg agactatgcc tccggagtcg gatgaggagg 840
acagcgggat cagcgatccc acggctttta tggctctctaa ggcgattgct atactaaaag 900
aagatgatga tggcgatgat ggatttcgac tggacgctcc ggcggtgccg gggagagact 960
gtaagtcctg tgaataccac cgggatcgta ccggagaccc gtctatgttg tgttctctgt 1020

```

gttatctccg	tcttaacgct	gcttttgtct	acagtaagtg	ttttgtgctt	ttttaccctg	1080
tggctttgtt	gagtttattt	ttttctgtgt	ctcatagggt	gttgtttatt	ataggtcctg	1140
tttcagatgt	ggaggaacct	gatagtacta	ctggaaatga	ggaggaaaag	ccctccccgc	1200
cgaaactaac	tcagcgctgc	agacctaata	ttttgagacc	ctcggcccag	cgtgtgtcat	1260
cccggaaacg	tgctgctgtt	aattgcatag	aagatttatt	ggaagagccc	actgaacctt	1320
tggacttgtc	cttaaagcga	ccccgcccgc	agtagggcgc	ggtgccagtt	ttttctctct	1380
agcttccggg	tgactcagtg	caataaaaaat	tttcttggca	acaggtgtat	gtgtttactt	1440
tacgggcggg	aagggttagg	gggagtataa	agctggaggg	gaaaaatctg	aggctgtcag	1500
atcgagttag	aagtccatg	gacttgtacg	agagcctaga	gaatctaagt	tctttgcgac	1560
gtttgctgga	ggaggcctcc	gacagaacct	cttacatttg	gaggtttctg	ttcggttccc	1620
ctctgagtcg	ctttttgcac	cgggtgaagc	gagagcacct	gacggaattt	gatgggcttt	1680
tagagcagct	gcctggactg	tttgattctt	tgaatctcgg	ccaccggacg	ctgctagagg	1740
agaggctttt	tccacaattg	gacttttcct	ctccaggccg	tctgtgttca	gcgcttgctt	1800
ttgctgtaca	tctgttggac	agatggaacg	agcagacgca	gctcagcccg	ggttacactc	1860
tggacttcct	gacgctatgc	ctatggaagt	tcggaatcag	gagggggagg	aagctgtacg	1920
ggcgcttggg	ggagaggcat	ccgtctctgc	gccagcagcg	tctgcaagct	caagtgtctg	1980
tgaggcggga	ggatctggaa	gccatttcgg	aggaggagag	cggcatggaa	gagaagaatc	2040
cgagagcggg	gctggaccct	ccggcggagg	agtagggggg	ataccggacc	cttttcctga	2100
gttggtcttg	ggggcggtgg	ggggcgcttc	tgtggtacgt	gaggatgaag	aggggcgcca	2160
acgcggtcag	aagaggggagc	attttgagtc	ctcgactttc	ttggctgatg	taaccgtggc	2220
cctgatggcg	aaaaacaggc	tggaggtggg	gtggtacccg	gaagtatggg	aggactttga	2280
gaaggggggac	ttgcacctgc	tggaaaaata	taactttgag	caggtgaaaa	catactggat	2340
gaacccggat	gaggactggg	aggtggtttt	gaaccgatac	ggcaaggtag	ctctgcgtcc	2400
cgactgtcgc	taccaggttc	gcgacaaggt	ggcctgcga	cgcaacgtgt	acctgttggg	2460
caacggcgcc	accgtggaga	tgggtgaccc	cagaaggggt	ggttttgtgg	ccaatatgca	2520
agaaatgtgc	cctgggggtgg	tgggcttgtc	tggggtgact	tttcatagt	tgaggtttag	2580
cggtagcaat	tttgggggtg	tggttattac	cgcgaaact	cctgtggtcc	tgcataattg	2640
ctactttttt	ggcttcagca	acacctgtgt	ggaaatgagg	gtgggaggca	aagtgcgcgg	2700
gtgttccttt	tacgcttgct	ggaagggggg	ggtgagccag	ggtaaggcta	aagtgtctgt	2760
tcacaagtgt	atgttggaga	gatgcacctt	gggcatttcc	agtgagggtc	tcctccacgc	2820
cagcgacaac	gtggcttctg	acaacggctg	cgcctttctt	atcaaggagg	ggggtcgcat	2880
ctgtcacaac	atgatatgcg	gccctgggga	tgtcccccca	aagccttacc	agatgggttac	2940
ctgcacagat	ggcaagggtg	gcatgctcaa	gcctgtgcac	attgtggggc	accggcgcca	3000
ccgctggcca	gagtttgaac	acaatgtgat	gacccgctgt	agcttgtacc	tgggaggcag	3060

gcgaggaggtt	ttcttgccca	gacagtgtaa	cctggcccac	tgcaacgtga	tcatggaaca	3120
atccgccgct	accaggttt	gctttggagg	aatatttgat	ataagcatgg	tggtgtataa	3180
gacctgctgc	tacgacgact	gtcgggctcg	tactcgaacc	tgcgactgcg	gagcctctca	3240
cctgtgtaac	ctgactgtga	tggggatggt	gactgaggag	gtgcgactgg	accactgtca	3300
gcactcttgc	ctgcgggagg	agttttcttc	ctcggacgag	gaggactagg	taggtggttg	3360
gggcgtggcc	agcgagaggg	tgggctataa	aggggaggtg	tcggctgacg	ctgtcttctg	3420
tttttcaggt	accatgagcg	gatcaagcag	ccagaccgcg	ctgagcttcg	acggggccgt	3480
gtacagcccc	tttctgacgg	ggcgcttgcc	tgccctgggg	ggagtgcgtc	agaatgttac	3540
cggttcgacc	gtggacggac	gtcccgtgga	tccatctaac	gctgcttcta	tgcgctacgc	3600
tactatcagc	acatctactc	tggacagcgc	cgctgccgcc	gcagccgcca	cctcagccgc	3660
tctctccgcc	gccaaagatca	tggctattaa	cccaagcctt	tacagccctg	tatccgtgga	3720
cacctcagcc	ctggagcttt	accggcgaga	tctagctcaa	gtggtggacc	aactcgagc	3780
cgtagagcaa	cagttgcagc	tggtgtcgac	ccgagtggag	caactttccc	gccctcccca	3840
gtaaccgcaa	aaattcaata	aacagaatth	aataaacagc	acttgagaaa	agtttaaact	3900
tgtggttgac	tttattcctg	gatagctggg	gggagggaac	ggcgggaacg	gtaagacctg	3960
gtccatcggt	cccggtcgtt	gagaacacgg	tggatttttt	ccaagacccg	atagaggtgg	4020
gtctgaacgt	tgagatacat	gggcatgagc	ccgtctcggg	ggtggaggta	ggccactgc	4080
agggcctcgt	tttcaggggt	ggtgttgtaa	atgatccagt	cgtaggcccc	ccgtgggcg	4140
tggtgctgga	agatgtcctt	cagcagcaag	ctgatggcaa	cgggaagacc	cttggtgtag	4200
gtgttgacaa	agcggttgag	ttgggagggg	tgcatgcggg	gactgatgag	gtgcattttg	4260
gcctggatct	tgaggttggc	tatgttgccg	cccagatcgc	gcctgggatt	catgttatgc	4320
aagaccacca	gcaccgagta	accggtgcag	cgggggaatt	tgctgtgcag	cttgggaagg	4380
aaagcgtgga	agaatttgga	gacccctcgg	tgccgccta	ggttttccat	gcactcatcc	4440
atgatgatgg	cgatgggccc	ccgggaggca	gcctgggcaa	aaacgttgcg	ggggtccgtg	4500
acatcgtagt	tgtggtcctg	ggtgagttca	tcataggaca	ttttgacaaa	gcgcgggcag	4560
agggtcccag	actggggaat	gatggttcca	tccggtccgg	gggcgtagtt	gccctcgag	4620
atttgcattt	cccaggcttt	gatttcagag	ggagggatca	tgtcaacctg	gggggcgatg	4680
aaaaaaatgg	tctctggggc	gggggtgatg	agctgggtgg	aaagcaggtt	gcgcaagagc	4740
tgtgacttgc	cgcagccggt	gggcccgtag	atgacagcta	tgacgggttg	cagggtgtag	4800
tttagagagc	tacaactgcc	atcatccttc	aaaagcgggg	ccacactgtt	taaaagttct	4860
ctaacatgta	agttttcccg	cactaagtcc	tgaggagac	gtgaccctcc	tagggagaga	4920
agttcaggaa	gcgaagcaaa	gtttttaagt	ggcttgaggc	catcggccaa	gggcaagtgc	4980
ctgagagttt	gactgagcag	ttccagccgg	tcccagagct	cggttacgtg	ctctacggca	5040
tctcgatcca	gcagacctcc	tcgtttcggg	ggttggggcg	gctctggctg	tagggaatga	5100
ggcgggtggc	gtccagctgg	gccatggtgc	ggtccctcca	tgggcgcagg	gttctcttca	5160

gggtggtctc	ggtcacggtg	aatgggtggg	ccccgggctg	ggcgctggcc	aggggtgcgct	5220
tgaggctgag	gcggctggtg	gcgaaccgtt	gcttttcgtc	tccctgcaag	tcagccaaat	5280
agcaacggac	catgagctca	tagtccaggc	tctctgcggc	atgtcctttg	gcgcgaagct	5340
tgccctttgga	aacgtgcccc	cagttttgagc	agagcaagca	ttttagcgcg	tagagttttg	5400
gcgccaagaa	cacggattcc	ggggaataag	catccccacc	gcagttggag	caaacggttt	5460
cgcattccac	cagccaggtc	agctgaggat	cttttgggtc	aaaaaccaag	cgcccgcctg	5520
tttttttgat	gcgcttccta	cctcgggtct	ccatgaggcg	gtgcccgcgt	tcggtgacga	5580
agaggctgtc	ggtgtctccg	tagacggagg	tcaggcgcg	ctcctccagg	ggggtcccgc	5640
ggtcctcggc	gtagagaaac	tcgcaccact	ctgacataaa	cgcccgggtc	caggctagga	5700
cgaatgaggc	gatgtgggaa	gggtaccggt	cggtatcgat	gaggggggtc	gtttttttcca	5760
aggtgtgcag	gcacatgtcc	ccctcgtccg	cttccaaaaa	tgtgattggc	ttgtaggtgt	5820
aagtcacgtg	atcctgtcct	tccgcggggg	tataaaaagg	ggcgtttccc	ccctcctcgt	5880
cactctcttc	cggttcgctg	tcgccaaagg	ccagctgttg	gggtacgtaa	acgcgggtga	5940
aggcgggcat	gacctgtgcg	ctgaggttgt	cagtttctat	atacgaggaa	gatttgatgg	6000
cgagcgcccc	cgtaggatg	cccttgagg	gctcggggcc	catttggtca	gaaaacacaa	6060
tctgtcgggt	atcaagcttg	gtggcaaaa	accgtagag	ggcgttggag	agcaacttgg	6120
cgatggagcg	ctgggttttg	ttttttttccc	ggtcggcttt	ttccttgggc	gcgatgttga	6180
gctggacgta	ctccctggcc	acgcacttcc	agccgggaaa	aacggccgtg	cgctcgtccg	6240
gcaccagcct	cacgtccat	ccgcgggtgt	gcagggtgat	gacgtcgatg	ctggtggcca	6300
cctctccgcg	caggggctcg	ttggtccagc	agaggcgacc	gcccttgcca	gagcagaagg	6360
ggggcagggg	gtcaagcagg	cgctcgtccg	gggggtcggc	gtcgatggta	aagatggcgg	6420
gcagcaggtg	tttgtcaaa	taatcgatct	gatgcccggg	gcaacgcagg	gcggttttccc	6480
agtcccgcac	cgccaaggcg	cgctcgtatg	gactgagggg	ggcgccccag	ggcatgggat	6540
gcgtcagggc	cgaggcgtag	atgccgcaga	tgtcatagac	gtaaaggggc	tcctccagga	6600
cgccgaggta	ggtggggtag	cagcgcccc	cgcggtatgt	ggcccgtacg	tagtcgtaga	6660
gctcgtgcga	ggggggcaga	aggtggcggc	tgaggtgagc	gcgctggggc	ttttcatctc	6720
ggaagaggat	ctgcctgaag	atggcggtgg	agttggagga	gatggtgggc	cgctgaaaaa	6780
tgttgaaagc	ggcgctcggc	agaccacagg	cctcgccgat	aaagtgggcg	taggactctt	6840
gcagcttttc	caccagggag	gcggtgacca	gcacgtccag	agcgcagtag	tccagggttt	6900
cccgcacgat	gtcataatgc	tcttcctttt	tttccttcca	gaggtctcgg	ttgaagagat	6960
actcttcgcg	gtctttccag	tactcttgga	gaggaaaccc	gttttcgtct	ccacggtaag	7020
agcccaacat	gtaaaactgg	ttgacggcct	gatagggaca	gcatcccttc	tccacgggca	7080
gcgagtaggc	cagggcgggc	ttgcgcaggg	aggtgtgagt	cagggcaaag	gtgtcgcgga	7140
ccataacttt	tacaaactgg	tacttaaagt	cccggtcgtc	gcacatgcct	cgctcccagt	7200

ctgagtagtc	tgtgcgcttt	ttgtgcttgg	ggtaggcag	ggagtaggtg	acgtcgttaa	7260
agaggatttt	gccacatctg	ggcataaagt	tgcgagagat	tctgaagggg	ccgggcacct	7320
ccgagcgggt	gttgatgact	tgggcagcca	ggagaatttc	gtcgaagccg	ttgatgttgt	7380
gccccacgac	gtagaactct	atgaaacgcg	gagcgccgcg	cagcaggggg	cacttttcaa	7440
gttgctggaa	agtaagtcc	cgcggtcga	cgccgtgttc	cgtgcggctc	cagtcctcca	7500
ccgggtttcg	ctccacaaaa	tcctgccaga	tgtggtcgac	tagcaagagc	tgagtcggt	7560
cgcgaaattc	gcggaatttt	ctgccgatgg	cttgcttctg	ggggttcaag	caaaaaaagg	7620
tgtctgcgtg	gtcgcgccag	gcgtcccagc	cgagctcgcg	agccagattc	agggccagca	7680
gcaccagagc	cggctcaccg	gtgattttca	tgacgaggag	aaagggcacc	agctgttttc	7740
cgaacgcgcc	catccaggtg	taggtctcca	cgctcgtaggt	gagaaacaga	cgttcgggtcc	7800
gcgggtgcga	tcccaggggg	aaaaacttga	tgggctgcca	ccattgggag	ctctgggcgt	7860
ggatgtgatg	gaagtaaaag	tcccggcggc	gcgtggaaca	ttcgtgctgg	tttttgtaaa	7920
agcggccgca	gtggtcgcag	cgcgagacgg	agtgaaggct	gtgaatcagg	tgaatcttgc	7980
gtcgtgagg	gggccccaga	gccaaaaagc	ggagcgggaa	cgaccgcgcg	gccacttcgg	8040
cgtccgcagg	caagatggat	gagggttcca	ccgttccccg	cccgcggacc	gaccagactt	8100
ccgccagctg	cggcttcagt	tcttgacca	gctctcgcag	cgtttcgtcg	ctgggcgaat	8160
cgtgaatacg	gaagttgtcg	ggtagaggcg	ggaggcgggtg	gacttccagg	aggtgtgtga	8220
gggccggcag	gagatgcagg	tgggtacttga	tttcccacgg	atgacggtcg	cgggcgtcca	8280
aggcgaagag	atgaccgtgg	ggccgcggcg	ccaccagcgt	tccgcggggg	gtctttatcg	8340
gcggcgggga	cgggctcccg	gcggcagcgg	cggctcggga	cccgcgggca	agtcgggcag	8400
cggcacgtcg	gcgtggagct	cgggcagggg	ctgggtgctgc	gcgcggagct	gactggcaaa	8460
ggctatcacc	cggcgattga	cgtcctggat	ccggcggcgc	tgctgaaga	ccaccggacc	8520
cgtggtcttg	aacctgaaag	agagttcgac	agaatcaatc	tcggcatcgt	taaccgcggc	8580
ctggcgcagg	atttcggcca	cgtccccgga	gttgtcttga	tacgcgattt	ctgccatgaa	8640
ctgggtcgatt	tcctcttcct	gcaagtctcc	gtgaccggcg	cgttcgacgg	tggccgcgag	8700
atcgttggag	atgcggccca	tgagctggga	aaaggcattg	atgccgacct	cgttccacac	8760
tcggctgtac	accacctctc	cgtgaacgtc	gcgggcgcgc	atcaccacct	gggcgagatt	8820
gagttccacg	tggcgggcga	aaaccggata	gtttcggagg	cgctgataca	gatagttgag	8880
gggtggtggc	gcgtgctcgg	ccacaaaaaa	atacatgatc	cagcggcgga	gggtcagctc	8940
gttgatgtcg	cccagcgctt	ccaggcggtt	catggcctcg	taaaagtcca	cggcaaagtt	9000
gaaaaattgg	ctgttcctgg	ccgagaccgt	gagctcttct	tccaagagcc	gaatgagatc	9060
cgccacgggtg	gccctgactt	cgcgttcgaa	agccccgggt	gcctcctcca	cctcttcctc	9120
ctcgacttct	tcgaccgctt	cgggcacctc	ctcttcctcg	accaccacct	caggcggggc	9180
tcggcggcgc	cggcggcgga	cgggcaggcg	gtcgacgaaa	cgctcgatca	tttccccctt	9240
ccgtcgacgc	atggtctcgg	tgacggcgcg	accctgttctg	cgaggacgca	gggtgaaggc	9300

gcccgcgccc agcggaggta acagggagat cggggggcg tctgggggga gactgacggc 9360
gctaactatg catctgatca atgtttgcgt agtgacctcg ggtcggagcg agctcagcgc 9420
ttgaaaatcc acgggatcgg aaaaccgttc caggaacgcg tctagccaat cacagtcgca 9480
aggtaagctg aggaccgtct cgggggcttg tctgttctgt cttcccgcgg tgggtgctgct 9540
gatgaggtag ttgaagtagg cgctcttgag gcggcggatg gtggacagga gaaccacgtc 9600
tttgcgccca gcttgctgta tccgcaggcg gtcggccatg ccccacactt ctccttgaca 9660
gcggcggagg tccttgtagt attcttgcat cagcctttcc acgggcacct cgtcttcttc 9720
ttccgctcgg ccggacgaga gccgcgtcag gccgtaccgg cgctgcccct gtggttgagg 9780
cagggccagg tcggccacga cgcgctcggc cagcacggcc tgctggatgc gggtaggggt 9840
gtcctgaaag tcgtcgagat ccacaaagcg gtggtacgcg ccagtgttga tgggtgtaggt 9900
gcagttgctc atgacggacc agtttacggg ctgggtgcca tggcccacgg tttccaggta 9960
gcggagacgc gagtagggcc gcgtctcgaa gatgtagtcg ttgcagggtcc gcagcaggta 10020
ctggtagccc accagcagat gcggcgggcg ctggcggtag aggggcccacc gctgggtggc 10080
gggggcgctg ggggcgagat cttccaacat gaggcggtga tagccgtaga tgtagcgcga 10140
catccaagtg atgccgctgg ccgtggtgct ggcgcggggc tagtcgcgaa cgcggttcca 10200
gatgtttcgc agcggctgga agtactcgat ggtggggcga ctctgccccg tgaggcgggc 10260
gcagtcggcg atgctctacg gggaaaaaga agggccagtg aacaaccggc ttccgtagcc 10320
ggaggagaac gcaagggggg caaagaccac cgaggctcgg gttcgaaacc cgggtggcgg 10380
cccgaatacg gagggcggtt ttttgctttt ttctcagatg catcccgtgc tgcggcagat 10440
gcgtccgaac gcgggggtccc agtccccggc ggtgcctgcg gccgtgacgg cggcttctac 10500
ggccacgtcg cgctccacc cgcctaccac ggcccaggcg gcggtggctc tgcgcggcgc 10560
aggggaaccc gaagcagagg cgggtgttga cgtggaggag ggccaggggt tggctcggct 10620
gggggcccct agtcccagac ggcacccgcg cgtggctctg aagcgcgacg cggcggaggc 10680
gtacgtgccg cggagcaatc tgtttcgcga ccgcagcggc gaggaggccg aggagatgcg 10740
agacttgctg tttcgggcgg ggaggaggt gcgtcacggg ctggaccggc agagggttct 10800
gagagaggag gactttgagg cggacgagcg cacgggggtg agtcccgcgc gggctcacgt 10860
ggcggccgcc aacctggtga gcgcgtacga gcagacggtc aaggaggaga tgaacttcca 10920
gaagagcttc aatcatcacg tgcgcacgct gattgcgcgc gaagaggtgg ccatcggcct 10980
catgcatctg tgggattttg tggaggcgta cggtcagaac cccagcagca agccgctgac 11040
ggctcagctg ttctcatcg tgcaacatag tcgagacaac gaaacgttca gggaggccat 11100
gctgaacatt gcagagcctg aggggcgctg gctcttgat ctcattaaca tcttgacagag 11160
tatcgtagtg caggagcgt cgctgagcct ggccgacaag gtggctgcca tcaactacag 11220
catgctgtcg ctgggcaaat tttagcccc caagatctac aagtctccgt tctccccat 11280
agacaaggag gtgaagatag acagctttta catgcgcatg gcgctcaagg tgctgactct 11340

aagcgacgac	ctgggggtgt	accgcaacga	ccgcatacac	aaggcggtga	gcgccagccg	11400
ccggcgcgag	ctgagcgacc	gcgagctttt	gcacagcctg	catcgggcgt	tgactggtgc	11460
cggcagcgcc	gaggcggccg	agtactttga	cgccggagcg	gacttgcgct	ggcagccatc	11520
ccgacgcgcg	ctggaggcgg	ctggcgtcgg	ggagtacggg	gtcgaggacg	acgatgaagc	11580
ggacgacgag	ttgggcattg	acttgtagcc	gtttttcggt	agatatgtcg	gcgaacgagc	11640
cgtctgcggc	cgccatggtg	acggcgggcg	gcgcgcccc	ggacccggcc	acgcgcgcgg	11700
cgtcgcagag	tcagccttcc	ggagtgcgcg	ccgcggacga	ctgggtccgag	gccatgcgtc	11760
gcacccctggc	gctgacggcg	cgcaaccccc	aggcttttcg	gcagcagccg	caggcaaacc	11820
ggtttgccgc	cattttggaa	gcggtggtgc	cctccagacc	caaccccacc	cacgaaaagg	11880
tgctggccat	cgtcaacgcc	ctggcgggaga	ccaaggccat	ccgccagac	gaggccgggc	11940
aggtttaca	cgcgctgcta	gaaagggtgg	gacgctacaa	cagctccaac	gtgcagacca	12000
atctggaccg	cttggtgacg	gacgtgaagg	aggccgtagc	ccagcgagag	cggtttttca	12060
aggaagccaa	tctgggctcg	ctggtggccc	tcaacgcctt	cctgagcacg	ctgccggcga	12120
acgtgccccg	cggtcaggag	gactacgtga	actttctgag	cgccctccgc	ctgatggtgg	12180
ccgaggtgcc	gcagagcgag	gtgtaccagt	ctggcccca	ctactacttc	cagacctccc	12240
ggcagggcct	gcagacggta	aacctgacgc	aggcctttca	gaacctgcag	ggccttttgg	12300
gggtgcgcgc	tccgctgggc	gaccgcagca	cgggtgtccag	cctgctgacc	cccaatgccc	12360
ggctgctctt	gcttctcatt	gctccgttca	ccgacagcgg	ttccatcagc	cgcgactctt	12420
acctgggaca	cctgctcacc	ctgtaccggg	aggccatcgg	gcaggcgcg	gtggacgagc	12480
agacgtacca	ggaaatcacc	agcgtgagcc	gcgcgctggg	gcaggaggac	acgggcagct	12540
tggaggcgac	tctgaacttc	ctgctgacca	accggcgcca	gcgcctacct	ccccagtagc	12600
cgtgaacgc	ggaggaggag	cgcatcctgc	gtttcgtgca	gcagagcacc	gcgctgtact	12660
tgatgcggga	aggcgctctt	cccagcgctt	cgttgacat	gacggcgcc	aacatggagc	12720
catcgttcta	cgccgccaac	cgtcccttcg	tcaaccggct	aatggactat	ttgcatcggg	12780
cggcgccct	gaacccggaa	tactttacta	acgtcatcct	gaacgaccgt	tggctgccac	12840
ctcccggctt	ctacacgggg	gagttcgacc	tcccggaggc	caacgacggt	ttcatgtggg	12900
acgacgtgga	cagcgtgttc	ctgcccggca	agaaggaggc	gggtgactct	cagagccacc	12960
gcgcgagcct	cgcagacctg	ggggcgaccg	ggcccgcgtc	tccgctgcct	cgcctgccga	13020
gcgccagcag	cgccagcgtg	gggcggggtga	gccgtccgcg	cctcagcggg	gaggaggact	13080
ggtggaacga	tccgctgctc	cgtccggccc	gcaacaaaa	cttccccaac	aacgggatag	13140
aggatttggt	agacaaaatg	aaccgttgga	agacgtatgc	ccaggagcat	cgggagtggc	13200
aggcgaggca	acccatgggc	cctgttctgc	cgccctctcg	gcgcccgcgc	agggacgaag	13260
acgccgacga	ttcagccgat	gacagcagcg	tggtggatct	gggcgggagc	gggaaccctt	13320
ttgcccacct	gcaacctcgc	ggcggtgggc	ggcggtggcg	ctaggaaaaa	aaattattaa	13380
aagcacttac	cagagccatg	gtaagaagag	caacaaaggt	gtgtcctgct	ttcttcccgg	13440

tagcaaaatg cgtcggggcgg tggcagttcc ctccgcggca atggcgtag gcccgcccc 13500
ttcttacgaa agcgtgatgg cagcggccac cctgcaagcg ccgttggaga atccttacgt 13560
gccgcccga tacctggagc ctacgggagg gagaaacagc attcgttact cggagctgac 13620
gcccctgtac gacaccaccc gcctgtacct ggtggacaac aagtcagcag atatcgccac 13680
cttgaactac cagaacgacc acagcaactt tctcacgtcc gtggtgcaga acagcgacta 13740
cacgcccgc gaagcgagca cgcagaccat taacttggac gaccgctcgc gctggggcgg 13800
ggacttgaaa accattctgc aactaacat gcccacgtg aacgagttca tgtttaccaa 13860
ctcgttcagg gctaaactta tgggtggcga cgaggccgac aaggaccggg tttatgagtg 13920
ggtgcagctg acgctgccgg aggggaactt ttcagagatt atgaccatag acctgatgaa 13980
caacgccatt atcgaccact acctggcggg agccagacag caggggggtga aagaaagcga 14040
gatcggcgtc aagtttgaca cgcgcaactt tcgtctgggc tgggaccggg agacggggct 14100
tgtgatgccg ggggtgtaca cgaacgaagc tttccatccc gacgtggtcc tcttgccggg 14160
ctgcgggggtg gactttacct acagccggtt aaacaacctg ctaggcatac gcaagagaat 14220
gccccttcag gaagggttcc agatcctgta cgaggacctg gagggcggtg acatcccggc 14280
cctgctggac gtgccggcgt acgaggagag catcgccaac gcaaggagg cggcgatcag 14340
gggcgataat ttcgcggcgc agccccaggc ggctccaacc ataaaaccg ttttggaaga 14400
ctccaaaggg cggagctaca acgtaatagc caacaccaac aacacggctt acaggagctg 14460
gtatctggct tataactacg gcgacccgga gaaggggggtt agggcctgga ccctgctcac 14520
cactccggac gtgacgtgcg gttcagagca ggtctactgg tcgctgcctg acatgtacgt 14580
ggaccctgtg acgtttcgt ccacgcagca agttagcaac taccagtggt tgggagcgga 14640
gcttatgccg attcacagca agagctttta caacgagcag gccgtctact cacagctcat 14700
tcgtcagacc accgccctaa cgcacgtttt caaccgctt cccgagaacc aaatcctagt 14760
gcgacctcca gcgcccacca tcaccaccgt cagcgagaac gtgcccgtc taaccgatca 14820
cgggacgtg cttttgcaga acagcatccg cggagttcag cgagttacca tcacggacgc 14880
ccgtcgtcgg acctgtccct acgtctacaa agccttggga atcgtggccc cgcgcgtcct 14940
gtcgagtcgc actttctaga tgtccatcct catctctccc agcaacaata ccggttgggg 15000
tctgggcgtg accaaaatgt acggaggcgc caaacgacgg tccccacaac atcccgtgcg 15060
agtgcgcggg cacttttagag ccccatgggg gtcgcacacg cgcggggcga ccggccgaac 15120
caccgtcgac gacgtgatcg atagcgtggg ggccgacgcc cgcaactacc agcccgtcgc 15180
atccacgggtg gacgaagtca tcgacggcgt ggtggccgac gccagggcct acgcccgcag 15240
aaagtctcgt ctgcgccgcc gccgttcgct aaagcgcccc acggccgcca tgaaagccgc 15300
tcgctctctg ctgcgtcgcg cacgtatcgt gggtcgccgc gccgccagac gcgcagccgc 15360
caacgccgcc gccggccgag tgcgccgccg ggccgccag caggccgccg ccgccatctc 15420
cagtctatcc gcccccgac gcgggaatgt gtactgggtc agggactcgg ccaccggcgt 15480

gcgagttccc gtgagaaccc gtcctcctcg tccctgaata aaaagttcta agcccaatcg 15540
 gtgttccggt gtgtgttcag ctggtcatga ccaaacgcaa gtttaaagag gagctgctgc 15600
 aagcgctggt ccccgaaatc tatgcgccgg cgccggacgt gaaaccgctg cgcgtagaac 15660
 gcgtgaagaa gcaggaaaag ctagagacaa aagaggaggc ggtggcggtg ggagacgggg 15720
 aggtggagtt tgtgctctcg ttcgctccgc gtcggcgagt gaattggaag gggcgcaagg 15780
 tgcaacgggt gctgctccc ggcacgggtg tgtctttcac cccgggtgaa aaatccgcct 15840
 ggaagggcat aaagcgctg tacgatgagg tgtacgggga cgaagacatt ctggagcagg 15900
 cgctggatag aagcggggag tttgcttacg gcaagaggc gaggacgggc gagatcgcca 15960
 tcccgtgga cacttccaa cccaccccca gtctgaaacc cgtgacgctg caacagggtgt 16020
 tgccggtgag cgtccctcg cgacgcggca taaaacgcga gggcggcgag ctgcagccca 16080
 ccatgcagct cctggttccc aagaggcaga aactagagga cgtactggac atgataaaaa 16140
 tggagcccga cgtgcagccc gatattaaaa tccgtcccat caaagaagtg gcgccgggaa 16200
 tgggcgtgca gaccgtggac atccagattc ccatgaccag cgccgcacag gcggtagagg 16260
 ccatgcagac cgacgtgggg atgatgacgg acctgcccgc agctgctgcc gccgtggcca 16320
 gcgccgcgac gcaaacggaa gccggcatgc agaccgacc gtggacggag gcgcccgtgc 16380
 agccggccag aagacgcgtc agacggacgt acggccccgt ttctggcata atgccggagt 16440
 acgcgtgca tccttccatc atccccacc ccggctaccg ggggcgcacc taccgtccgc 16500
 gacgcagcac cactcgcgc cgtcgcgcga cggcacgagt cgccaccgcc agagtgcagac 16560
 gcgtaacgac acgtcgcggc cgccgcttga ccctgcccgt ggtgcgtac catcccagca 16620
 ttctttaaaa aaccgctcct acgttgacga tgggcaagct tacttgctga ctccgtatgg 16680
 ccgtgcccgg ctaccgagga agatcccgcc gacgacggac tttgggaggc agcggtttgc 16740
 gccgccgtcg ggcggttcac cggcgctca agggaggcat tctgccggcc ctgatcccca 16800
 taatcgccgc agccatcggg gccattcccg gaatcgccag cgtagcgggt caggctagcc 16860
 agcgcactg attttactaa ccctgtcggg cgcgccgtct ctttcggcag actcaacgcc 16920
 cagcatggaa gacatcaatt tctcctctct ggccccgcgg cagggcacgc ggccgtatat 16980
 ggggacgtgg agcgagatcg gcacgaacca gatgaacggg ggcgctttca attggagcgg 17040
 tgtgtggagc ggcttgaaaa atttcggttc cactctgaaa acttacggca accgggtgtg 17100
 gaactccagc acggggcaga tgctgagggg caagctaaag gacacgcagt ttcagcaaaa 17160
 ggtggtggac ggcacgctt cgggcctcaa cggcgccgtc gacctggcca accaggccat 17220
 tcaaaaggaa attaacagcc gcctggagcc gcggccgcag gtggaggaga acctgcccc 17280
 tctggaggcg ctgcccccca agggagagaa gcgccgcgg cccgacatgg aggagacgct 17340
 agttactaag agcgaggagc cgccatcata cgaggaggcg gtgggtagct cgagctgcc 17400
 gtccctcacg ctgaagccca ccacctatcc catgaccaag cccatcgctt ccatggcgcg 17460
 ccccggtgga gtcgacccgc ccatcgacgc ggtggccact ttggacctgc cgcgccccga 17520
 acccggcaac cgcgtgcctc ccgtcccat cgctccgcg gtttctcgcc ccgcatccg 17580

ccccgtcgcc gtggccactc cccgctatcc gagccgcaac gccaaactggc agaccaccct 17640
caacagtatt gtcggactgg ggggtgaagtc tctgaagcgc cgtcgctgtt tttaaagcac 17700
aattttattaa acgagtagcc ctgtcttaat ccatcgttgt atgtgtgcct atatcacgcg 17760
ttcagagcct gaccgtccgt caagatggcc actccgtcga tgatgcegca gtggtcgtac 17820
atgcacatcg ccgggcagga cgccctcgag tacctgagcc cgggtctggg gcagtttgcc 17880
cgtgcgacgg aaacctactt ctactgggc aacaagttca ggaacccac cgtggcgccc 17940
accacgacg tcaccaccga tcggtcccag cgactgacaa tccgcttcgt ccccgctggac 18000
aaggaagaca ccgcttactc ctacaaaacc cgcttcacgc tggccgtggg cgacaaccgg 18060
gtgctagaca tggccagtac ctactttgac atccgcggcg tgatcgaccg cggacctagc 18120
ttcaagcctt actccggcac ggcttacaac tctactggctc ccaaaggggc gcccaacaac 18180
agccaatgga acgccacaga taacgggaac aagccagtgt gttttgctca ggcagctttt 18240
ataggtcaaa gcattacaaa agacggagtg caaatacaga actcagaaaa tcaacaggct 18300
gctgccgaca aaacttacca accagagcct caaattggag tttccacctg ggataccaac 18360
gttaccagta acgctgccgg acgagtgtta aaagccacca ctcccatgct gccatgttac 18420
ggttcatatg ccaatccac taatccaaac gggggtcagg caaaaacaga aggagacatt 18480
tcgctaaact ttttcacaac aactgcggca gcagacaata atcccaaagt ggttctttac 18540
agcgaagatg taaaccttca agccccgat actcacttag tatataagcc aacggtggga 18600
gaaaacgtta tcgccgcaga agccctgcta acgcagcagg cgtgtcccaa cagagcaaac 18660
tacatagggt tccgagataa ctttatcggt ttaatgtatt ataacagcac agggaacatg 18720
ggagtctctg caggtcaggc ctcgcagtta aacgcagttg tagacctgca agatcgaaac 18780
acggaactgt cctatcagct aatgctagat gctctgggtg acagaactcg atatttctca 18840
atgtggaatc aggccgtgga cagctacgat ccagacgtta ggattatcga gaacctggg 18900
gtggaagacg agctgccc aa ttactgtttt ccactcccag gcatgggtat ttttaactcc 18960
tacaaggggg taaaaccaca aaatggcggg aatggtaact gggaagcaaa cggggacctt 19020
tcaaagtcca atgagatcgc tttaggaaac atttttgcca tggaaattaa cctccacgca 19080
aacctgtggc gcagcttctt gtacagcaat gtggcgctgt acctgccaga cagctataaa 19140
ttcactccc ctaacatcac tctgcccgcc aaccaaaaca cctacgagta tatcaacggg 19200
cgcgtcactt ctccaacct ggtggacacc tttgttaaca ttggagccc atggtcgcg 19260
gatcccatgg acaacgtcaa cccctttaac catcaccgga acgcgggcct ccgttaccgc 19320
tccatgctgc tgggaaatgg acgcgtgggt cttttccaca tacaagtgcc gcaaaaattt 19380
ttcgcgatta agaacctctt gcttttgccc ggctcctaca cttacgagt gagcttcaga 19440
aaagacgtga acatgattct gcagagcacc ctgggcaatg atcttcgagt ggacggggcc 19500
agcgtccgca ttgacagcgt caacttgtag gccaaacttt tccccatggc gcacaacacc 19560
gcttctacct tggaagccat gctgcgaaac gacaccaacg accagtcgtt taacgactac 19620

ctcagcgcg ccaacatgct ttatcccatt ccggccaacg ccaccaacgt tcccatttcc 19680
 attccctccc gcaactgggc ggccttccgg ggatggagct tcacccgcct taaagccaag 19740
 gaaacgcctt ccttgggctc cggttttgac ccctactttg tgtactcagg caccattcct 19800
 tacctggacg gcagctttta cctcaaccac actttcaaac gtctgtccat cgtgttcgat 19860
 tcttccgtaa gttggccggg caacgaccgc ctctgacgc cgaacgagtt cgaaattaag 19920
 cgcatgtgg acggggaagg ctacaacgtg gctcaaagta acatgaccaa agactgggtt 19980
 ttaattcaaa tgctcagcca ctacaacatc ggctaccaag gcttctatgt tcccgaagggc 20040
 tacaaggatc ggatgtattc tttcttccga aactttcagc ccatgagccg ccagggtgccg 20100
 gatcccaccg ctgccggcta tcaagccgtt cccctgccca gacaacacaa caactcgggc 20160
 tttgtggggg acatggggcc gaccatgcgc gaaggacagc catacccggc caactacccc 20220
 tatcccctga tcggcgctac cgccgtcccc gccattacc agaaaaagtt tttgtgcgac 20280
 cgcgatcatgt ggcgcatacc tttttccagc aactttatgt caatggggggc cctgaccgac 20340
 ctcgacaga acatgcttta cgctaactcc gcccatgccc tggatatgac ttttgagggtg 20400
 gaccccatga acgagcccac gttgctgtac atgctttttg aggtgttcga cgtgggtcaga 20460
 gtgcaccagc cgcaccgcgg tattatcgag gccgtgtacc tgcgcacccc cttctctgcg 20520
 ggcaatgcca ccacataagc cgctgaacta gctgggtttt accccagatc ccatgggctc 20580
 cacggaagac gaactgcggg ccattgtgcg agacctgggc tgcggaccct acttcctggg 20640
 cacctttgac aagcggtttc ccgggttcgt gtctcctcgc aaactcgcgt gcgcatcgt 20700
 gaataccgcc ggccgagaga ccggaggaga gcattggcta gctctgggct ggaacccccg 20760
 ctctccacg tttttcctgt tcgaccctt tggcttttca gaccaacgct tgaagcagat 20820
 ctatgcattt gaatatgagg gtctactcaa gcgaagcgcg ctggcctcct ccgccgatca 20880
 ctgtctaacc ctggtaaaga gactcagac ggttcagggc cctcacagcg ccgcctgtgg 20940
 ctttttttgt tgcattgttt tgacgcctt tgtgaactgg ccggacaccc ccatggaaaa 21000
 caacccacc atggacctcc tgactggcgt tcccaactcc atgctccaaa gcccagcgt 21060
 gcagaccacc ctctccaaa accagaaaaa tctgtacgcc tttctgcaca agcactctcc 21120
 ctactttcgc cgccatcggg aacaaataga aaatgcaacc gcgtttaaca aaactctgta 21180
 acgtttaata aatgaacttt ttattgaact ggaaaacggg tttgtgattt ttaaaaatca 21240
 aaggggttga gctggacatc catgtgggag gccggaaggg tgggtgttct gtactggtac 21300
 ttgggcagcc acttaaaactc tggaaatcaca aacttgggca gcggtatttc tgggaagtgtg 21360
 tcgtgccaca gctggcgggt cagctgaagt gcctgcagaa catcgggggc ggagatcttg 21420
 aagtcgcagt ttatctgggt cacggcacgc gcgttgcggt acatgggatt ggcacactga 21480
 aacaccagca ggctgggatt cttgatgcta gccagggcca cggcgtcgggt cacgtcaccg 21540
 gtgtcttcta tgttggacag cgaaaaaggc gtgactttgc aaagctggcg tcccgcgcga 21600
 ggcacgcaat ctcccaggta gttgactca cagcggatgg gcagaagaag atgcttgtgg 21660
 ccgcggttca tgtagggata ggccgctgcc ataaaagctt cgatctgcct gaaagcctgc 21720

ttggccttgt gcccttcggt ataaaaaaca ccgcaggact tgttggaataa ggtattactg 21780
gcgcaagcgg catcgtgaaa gcaagcgcgt gcgtcttcgt ttcgtaactg caccacgctg 21840
cggccccacc ggttctgaat caccttgggc ctgccgggggt tttccttgag agcgcgctgg 21900
ccggcttcgc tgcccacatc catttccacg acatgctcct tgtaaatcat ggccagaccg 21960
tggaggcagc gcagctcctc gtcacgtcgc gtgcagtgat gctccacac gacgcagcca 22020
gtgggctccc acttgggctt ggaggcctcg gcaatgccag aatacaggag aacgtagtgg 22080
tgcagaaaac gtcccatcat ggtgccaaag gttttctggc tgctgaaggc catcgggcag 22140
tacctccagt cctcgtaag ccaagtgtt cagatcttcc tgaagaccgt gtactgatcg 22200
ggcataaagt ggaactcatt gcgctcggtc ttgtcgatct tatacttttc catcagacta 22260
tgcataatct ccatgccctt ttcccaggcg caaacaatct tgggtgctaca cgggttaggt 22320
atggccaaag tggttggcct ctgaggcggc gcttgttctt cctcttgagc cctctcccga 22380
ctgacggggg ttgaaagagg gtgccccttg ggaacggct tgaacacggc ctggcccag 22440
gcgtcccgaa gaatctgcat cgggggattg ctggccgtca tggcgatgat ctgaccccgg 22500
ggctcctcca ctctgtcctc ctgaggactt tcctcgtgct tttcggggga cggtagcgga 22560
gtagggggaa gagcgcggcg cgccttcttc ttgggaggca gttccggagc ctgctcttga 22620
cgactggcca ttgtcttctc ctaggcaaga aaaacaagat ggaagactct ttctcctcct 22680
cctcgtcaac gtcagaaagc gagtcttcca ccttaagcgc cgagaactcc cagcgcatag 22740
aatccgatgt gggctacgag actccccccg cgaacttttc gccgcccccc ataaacacta 22800
acgggtggac ggactacctg gccctaggag acgtactgct gaagcacatc aggcggcaga 22860
gcgttatcgt gcaagatgct ctaccgagc gactcgcggc tccgctggaa gtggcggaac 22920
ttagcgcgc ctacgagcga accctcttct ccccaaagac tcccccaag aggcaggcta 22980
acggcacctg cgagcctaac cctcgactca acttctaccc tgcctttgcc gtgccagagg 23040
tactggctac gtaccacatt tttttccaaa accacaaaat ccctctctcg tgccgcgcca 23100
accgcacca agccgatcgc gtgctgcgac tggaggaagg ggctcgcata cctgagattg 23160
cgtgtctgga ggaagtccca aaaatctttg aaggctctggg ccgcgacgaa aagcgagcag 23220
caaacgctct ggaagagaac gcagagagtc acaacagcgc cttggtagaa ctcgagggcg 23280
acaacgccag actggccgct ctcaaacggt ccatagaagt cacgcacttc gcctaccccg 23340
ccgttaacct ccctccaaaa gttatgacag cggcatgga ctcgctgctc ataaagcgcg 23400
ctcagccctt agaccagag cacgaaaaca acagtacga aggaaaaccg gtggtttctg 23460
atgaggagt gagcaagtgg ctgtcctcca acgacccgc cacgttgag gaacgaagaa 23520
aaaccatgat ggccgtggtg ctagttaccg tgcaattaga atgtctgcag aggttctttt 23580
cccaccaga gaccctgaga aaagtggagg aaacgctgca ctacacattt aggcacggct 23640
acgtgaagca agcctgcaag atttccaacg tagaacttag caacctcatc tcctacctgg 23700
ggatcttgca cgaaaaccgc ctcggaacaa acgtgctgca cagcactg aaaggagaag 23760

cccgccgaga ctatgtgcga gactgcggtg tcctagcgct agtgtacacc tggcagagcg 23820
 gaatgggagt ctggcagcag tgcctggagg acgaaaacct caaagagctt gaaaagctgc 23880
 tgggtgcgctc cagaagggca ctgtggacca gttttgacga gcgcaccgcc gcgcgagacc 23940
 tagctgatat ttttttctt cccaagctgg tgcagactct ccgggaagga ctgccagatt 24000
 ttatgagtca aagcatcttg caaaacttcc gctctttcat cttggaacgc tcgggaatct 24060
 tgcccgccac tagctgcgcc ctaccacacag attttgtgcc tctccactac cgcgaaatgcc 24120
 caccgccgct gtggccgtac acttacttgc ttaaactggc caactttcta atgttccact 24180
 ctgacctggc agaagacgtt agcggcgagg ggctgctaga atgccactgc cgctgcaacc 24240
 tgtgcacccc ccaccgtct ctagtatgca acactcccct gctcaatgag acccagatca 24300
 tcggtacctt tgaaatccag ggaccctccg acgcggaaaa cggcaagcag gggctctgggc 24360
 taaaactcac agccggactg tggacctccg cctacttgcg caaatttgta ccagaagact 24420
 atcacgccc ccaaattaaa ttttacgaaa accaatcaaa accacccaaa agcgagttaa 24480
 cggcttgctg cattacgcag agcagcatag ttgggcagtt gcaagccatt aacaaagcgc 24540
 ggcaagagtt tctcctaaaa aaaggaaaag gggcttactt ggacccccag accggcgagg 24600
 aactcaacgg accctctca gtcgcaggtt gtgtgcccc tgcgcgccc aaagaacacc 24660
 tcgcagtgga acatgccaga gacggaggaa gaggagtgga gcagtgtgag caacagcgaa 24720
 acggaggaag agccgtggcc cgaggggtgc aacggggaag aggacacgga gggacggcga 24780
 agtcttcgcc gaagaactct cgccgctgcc cccgaagtcc cagccggccg cctcggccca 24840
 agatcccgca cacaccgta gatgggatag caagaccaa aagccgggta agagaaacgc 24900
 tcgccccgc cagggctacc gctcgtggag aaagcaciaa aactgcatct tatcgtgctt 24960
 gctccagtgc ggcgagacg tttcgttcac ccgtagatac ttgcttttta acaaaggggt 25020
 ggccgtcccc cgtaacgtcc tccactacta ccgtcactct tacagctccg aagcggacgg 25080
 ctaagaaaac gcagcagttg ccggcgggag gactgcgctc cagcgcccga gaacccccag 25140
 ccaccaggga gtcgcgaac cgcatatttc ccaccctcta cgctatcttt cagcaaagcc 25200
 gggggcagca gcaagaactg aaaataaaaa accgcacgct gaggtcgctt acccgaagct 25260
 gcctctatca caagagcgaa gagcagctgc agcgaaccct ggaggacgca gaagcgctgt 25320
 tccagaagta ctgcgcgacc accctaaata actaaaaaag ccgcgcgcgc ggacttcaaa 25380
 ccgtctgacg tcaccagccg cgcgcacaaa tgagcaaaga gattcccacg ccttacatgt 25440
 ggagttacca gccgcagatg ggattagccg ccggcgcgcg ccaggattac tccacgaaaa 25500
 tgaactggct cagcgccggg cccacatga tttcccgcgt aaacgacatt cgcgcccacc 25560
 gcaatcagct attgttagaa caggctgctc tgaccgccac gcccgcgtat aacctgaacc 25620
 ctcccagctg gccagctgcc ctggtgtacc aggaaacgcc tccaccacc agcgtacttt 25680
 tgccccgtga cggccaggcg gaagtccaga tgactaacgc gggcgcgcaa ttagcgggcg 25740
 gatcccggtt tcggtacaga gttcacggcg ccgcacccta tagcccaggt ataaagaggc 25800
 tgatcattcg aggcagaggt gtccagctca acgacgagac agtgagctct tcgcttggtc 25860

tacgaccaga cggagtgttc cagctcgcgg gctcggggccg ctcttcgttc acgcctcgcc 25920
aggcatacct gactctgcag agctctgcct ctcagcctcg ctcgggagga atcggacccc 25980
ttcagtttgt ggaggagttt gtgccctcgg tctactttca gcctttctcc ggatcgcccg 26040
gccagtaccg ggacgagttc atccccaact tcgacgcggg gagtgactct gtggacgggt 26100
atgactgatg tcgagcccg cttcagtgcta gtggaacaag cgcggctcaa tcacctgggt 26160
cgttgccgcc gccgctgctg cgtggctcgc gacttgagct tagctctcaa gtttgtaaaa 26220
aaccgcgtccg aaaccgggag cgctgtgcac gggttggagc tagtgggtcc tgagaaggcc 26280
accatccacg ttctcagaaa ctttgtggaa aaaccattt tggttaaacg agatcagggg 26340
ccttttgtaa tcagcttact ctgcacctgt aaccatgttg accttcacga ctattttatg 26400
gatcatttgt gcgctgaatt caataagtaa agcgaattct taccaagatt atgatgtcca 26460
tgactgttcc tcgccactat acgatgttgt gccagtaaac tctcttgctg acatctatct 26520
gaactgttcc ttttggtccg cacagcttac ttggtactac ggtgacaccg tcctttctgg 26580
ctcactgggc agctcacacg gaataacact tcacctctt tcgccgtttc gatacggaaa 26640
ctacagctgt cgtgccggtta cctgcctcca cgttttcaat cttcagccct gtccaccgac 26700
caaacttgta tttgtcgact ctaagcactt acagctcaac tgcagcattc taggccccag 26760
tatcttggtg acatacaata aaatcagggt ggtggaattt gtctactacc caccagcgc 26820
ccgcggtttt ggggaaattc ctttcagat ctactacaac tatcttgcca cacattatgc 26880
aagtcaacag caactaaact tgcaagcacc cttcacgcca ggagagtact cctgtcacgt 26940
aggctcctgc acagaaactt ttattctctt caacagatct tctgccattg aacgcttcac 27000
tactaactac tttagaaacc aagttgtgct tttcactgac gaaacccta acgtcacct 27060
ggactgtgca tgtttttctc atgacaccgt aacttggact cttacaata ctctctggct 27120
cgcgttcgat aaccaaagct tgattgttaa aaattttgat ttaaccttta ctaaaccctc 27180
tcctcgcgaa atagttatct ttgtccttt taatccaaaa actaccttag cctgtcaggt 27240
tttgtttaag cttgccaaa caaactttaa gtttgttat ttgcctccgc aatctgtcaa 27300
actcatagaa aaatacaaca aagcggcgt cttggctcct aaaaccttct accactggct 27360
aacctacacg gggctgttg cactaattgt ttttttcta attaacattt ttatatgttt 27420
cttgccctcc tccttcttt cgcgaaacac gttgccgcag aaagacctct ccttattact 27480
gtagcgcttg ctatacaaaa ccaagagtgg tcaaccgtgc tctcaatcta ttttcaattt 27540
ttcattttgt ccttaatact ttctcttatt gtcgttaaca atgatctgga gcattggtct 27600
cgcctttttt tggctgctta gtgcaaaagc cactattttt cacaggatg tggagaagg 27660
aactagcacc ctctttacga tacctgaaac aattaaggcg gctgatgaag tttcttggtta 27720
caaaggctcg ctctcagacg gcaaccactc attctcagga cagaccctt gcacccaaga 27780
aacttatttt aaatcagaac tacaatacag ctgcataaaa aactttttcc atctctacaa 27840
catctcaaaa ccctatgagg gtattttacaa tgccaagggt tcagacaact ccagcacacg 27900

gaactttttac	tttaaatctga	cagttatttaa	agcaattttcc	attcctatct	gtgagtttag	27960
ctcccagttt	ctttctgaaa	cctactgttt	aattactata	aactgcacta	aaaatcgcct	28020
tcacaccacc	ataatctaca	atcacacaca	atcaccttgg	gttttaaacc	taaaattttc	28080
tccacacatg	ccttcgcaat	ttctcacgca	agttaccgtc	tctaacataa	gcaagcagtt	28140
tggctttttac	tatccttttc	acgaactgtg	cgaaataatt	gaagccgaat	atgaaccaga	28200
ctactttact	tacattgcca	ttgggtgtaat	cgttgtttgc	ctttgctttg	ttattggggg	28260
gtgtgtttat	ttgtacattc	agagaaaaat	attgctctcg	ctgtgctcct	gcggttacaa	28320
agcagaagaa	agaattaaaa	tctctacact	ttattaatgt	tttccagaaa	tggcaaaact	28380
aacgctccta	cttttgcttc	tcacgccggt	gacgcttttt	accatcactt	tttctgccgc	28440
cgccacactc	gaacctcaat	gtttgccacc	ggttgaagtc	tactttgtct	acgtgttgct	28500
gtgctgcgtt	agcgtttgca	gtataacatg	ttttaccttt	gtttttcttc	agtgcattga	28560
ctactttctg	gtcagactct	actaccgcag	acacgcgcct	cagtatcaaa	atcaacaaat	28620
tgccagacta	ctcggctctg	catgattgtc	ttgtatttta	ccctgatttt	ttttcacctt	28680
acttgcgctt	gtgattttca	cttcaactca	ttttggaaaa	cgcaatgctt	cgacccgcgc	28740
ctctccaacg	actggatgat	ggctcttgca	attgccacgc	ttggggcggt	tggacttttt	28800
agtgggtttg	ctttgcatta	caaatttaag	actccatgga	cacatggctt	tctttcagat	28860
tttccagtta	cacctactcc	gccgcctccc	ccggccatcg	acgtgcctca	ggttccctca	28920
ccttctccat	ctgtctgcag	ctactttcat	ctgtaatggc	cgacctagaa	tttgacggag	28980
tgcaatctga	gcaaaggggt	atacacttcc	aacgccagtc	ggaccgcgaa	cgcaaaaaca	29040
gagagctgca	aaccatacaa	aacaccacc	aatgtaaacg	cgggatatatt	tgtattgtaa	29100
aacaagctaa	gctccactac	gagcttctat	ctggcaacga	ccacgagctc	caatacgtgg	29160
tcgatcagca	gcgtcaaacc	tgtgtattct	taattggagt	ttccccatt	aaagtactc	29220
aaaccaaggg	tgaaccaag	ggaaccataa	ggtgctcatg	tcacctgtca	gaatgccttt	29280
acactctagt	taaaacccta	tgtggcttac	atgattctat	cccctttaat	taaataaact	29340
tactttaaat	ctgcaatcac	ttcttcgtcc	ttgtttttgt	cgccatccag	cagcaccacc	29400
ttcccctctt	cccaactttc	atagcatatt	ttccgaaaag	aggcgtactt	tcgccacacc	29460
ttaaagggaa	cgttttacttc	gctttcaagc	tctcccacga	ttttcattgc	agatatgaaa	29520
cgcgccaaag	tggaagaagg	atttaacccc	gtttatccct	atggatattc	tactccgact	29580
gacgtggctc	ctccctttgt	agcctctgac	ggtcttcaag	aaaaccacc	tggggctctg	29640
tccctaaaaa	tatccaaacc	tttaactttt	aatgcctcca	aggctctaag	cctggctatt	29700
ggtccaggat	taaaaattca	agatggtaaa	ctagtggggg	agggacaagc	aattcttgca	29760
aacctgccgc	ttcaaatac	caacaacaca	atttactac	gttttgggaa	cacacttgcc	29820
ttgaatgaca	ataatgaact	ccaaaccaca	ctaaaatctt	catcgcccct	taaaatcaca	29880
gaccagactc	tgtcccttaa	catagggggac	agccttgcaa	ttaaagatga	caaactagaa	29940
agcgtctctc	aagcgaccct	cccactctcc	attagcaaca	acaccatcag	cctcaacgtg	30000

ggcaccggac tcaccataaa tggaaacggt ttacaagctg ttcccttaaa tgctctaagt 30060
cccctaacta tttccaacaa taacatcagc ctgcgctatg gcagttccct gacggtgctt 30120
aacaatgaac tgcaaagcaa cctcacagtt cactcccctt taaaactcaa ctccaacaac 30180
tcaattttctc tcaacactct atctccgttt agaatcgaga atggtttcct cacgctctat 30240
ttgggaacaa aatctggctt gctagttcaa aacagtggct taaaagtcca agcgggctac 30300
ggcctgcaag taacagacac caatgctctc acattaagat atctcgctcc actgaccatt 30360
ccagactcgg gctcagaaca aggcattctt aaagtaaaca ctggacaggg cctaagtgtg 30420
aaccaagctg gagcgcttga aacatcccta ggaggtggat taaaatatgc tgataacaaa 30480
ataacctttg atacaggaaa cggactgaca ttatctgaaa ataaacttgc agtagctgca 30540
ggtagtggtc taacttttag agatgggtgcc ttggtagcca cggaaccgc atttacgcaa 30600
acactgtgga ctacggctga tccgtctccc aactgcacaa ttatacagga ccgcgacaca 30660
aaatttactt tggcgcttac cattagtggg agccaagtgc tggggacggt ttccattatt 30720
ggagtaaaaag gccccctttc aagtagcata ccgtcagcta ccgttacagt acaacttaac 30780
tttgattcca acggagccct attgagctcc tcttcactta aaggttactg ggggtatcgc 30840
caaggccct caattgacct ttacccata attaatgcct taaactttat gccaaactca 30900
ctggcttatc ccccgggaca agaaatccaa gcaaaatgta acatgtacgt ttctactttt 30960
ttacgaggaa atccacaaag accaatagtt ttaaacatca cttttaataa tcaaaccagc 31020
gggttttcca ttagatttac atggacaaat ttaaccacag gagaagcatt tgcaatgccc 31080
ccatgcactt tttctacat tgctgaacaa caataaacta tgtaaccctc accgttaacc 31140
cgcctccgcc ctccattttt attttataaa ccacccgatc caccttttca gcagtaaaca 31200
attgcatgtc agtaggggca gtaaaacttt tgggagttaa aatccacaca ggttcttcac 31260
aagctaagcg aaaatcagtt acacttataa aaccatcgct aacatcggac aaagacaagc 31320
atgagtccaa agcttccggt tctggatcag atttttgttc attaacagcg ggagaaacag 31380
cttctggagg attttccatc tccatctcct tcatcagttc caccatgtcc accgttgtca 31440
tctgggacga gaacgacagt tgtcatacac ctcataagtc accggtcgat gacgaacgta 31500
cagatctcga agaatgtcct gtcgccgctt ttcggcagca ctgggccgaa ggcgaaagcg 31560
cccatgttta acaatggcca gcaccgccc cttcatcagg cgcctagtcc ttttagcgca 31620
acagcgcatg cgcagctcgc taagactggc gcaagaaaca cagcacagaa ccaccagatt 31680
gttcatgatc ccataagcgt gctgacacca gcccatacta acaaattggt tcactattct 31740
agcatgaatg tcatatctga tgttcaagta aattaaatgg cgccccctta tgtaaact 31800
tcccacgtac aacacctcct ttggcatctg ataattaacc acctccgat accaaatata 31860
tctctgatta atagtcgccc cgtacactac ccgattaaac caagttgcca acataatccc 31920
ccctgccata cactgcaaag aacctggacg gctacaatga cagtgcaaag tccacacctc 31980
gttgccatgg ataactgagg aacgccttaa gtcaatagtg gcacaactaa taaaacatg 32040

taaatagtgt	ttcaacaagt	gccactcgta	tgaggtgagt	atcatgtccc	agggaaacggg	32100
ccactccata	aacactgcaa	aaccaacaca	tcctaccatc	ccccgcacgg	cactcacatc	32160
gtgcatggtg	ttcatatcac	agtccggaag	ctgaggacaa	ggaaaagtct	cgggagcatt	32220
ttcatagggc	ggtagtgggt	actccttgta	ggggttcagt	cggcacagggt	atctcctcac	32280
cttctgggcc	ataacacaca	agttgagatc	tgatttcaag	gtactttctg	aatgaaaacc	32340
aagtgccttc	ccaacaatgt	atccgatgtc	ttcgggtccc	gcgtcggtag	cgctccttgc	32400
agtacacacg	gaacaaccac	tcacgcaggc	ccagaagaca	gttttccgcg	gacgggtgaca	32460
agttaatccc	cctcagtctc	agagccaata	tagtttcttc	cacagtagca	taggccaaac	32520
ccaaccagga	aacacaagct	ggcacgtccc	gttcaacggg	aggacaagga	agcagaggca	32580
gaggcatagg	caaagcaaca	gaatttttat	tccaactggt	cacgtagcac	ttcaaacacc	32640
aggtcacgta	aatggcagcg	atcttgggtt	tcctgatgga	acataacagc	aagatcaaac	32700
atgagacgat	tctcaagggt	attaaccaca	gctggaatta	aatcctccac	gcgcacattt	32760
agaaacacca	gcaatacaaa	agcccgggtt	tctccgggat	ctatcatagc	agcacagtca	32820
tcaattagtc	ccaagtaatt	ttcccgtttc	caatctgtta	taatttgtag	aataatgccc	32880
tgtaaatacca	agccggccat	ggcgaaaagc	tcagataatg	cactttccac	gtgcattcgt	32940
aaacacaccc	tcattctgtc	aatccaaaaa	gtcttcttct	tgagaaacct	gtagtaaatt	33000
aagaatcgcc	aggttagggt	cgatgcctac	atcccggagc	ttcatttctca	gcatgcactg	33060
caaatgatcc	agcagatcag	aacagcaatt	agcagccagc	tcattccccg	tttccagttc	33120
cggagttccc	acggcaatta	tcactcgaaa	cgtgggacaa	atcgaaataa	catgagctcc	33180
cacgtgagca	aaagccgtag	ggccagtgcg	ataatcacag	aaccagcgga	aaaaagattg	33240
cagctcatgt	ttcaaaaagc	tctgcagatc	aaaattcagc	tcattgcaat	aacacagtaa	33300
agtttgcggt	atagtaaccg	aaaaccacac	gggtcgacgt	tcaaactctt	cggcttacct	33360
aaaaaagaag	cacattttta	aaccacagtc	gcttcttgaa	caggaggaaa	tatggtgcgg	33420
cgtaaaacca	gacgcgccac	cggatctccg	gcagagccct	gataatacag	ccagctgtgg	33480
ttaaacagca	aaacctttta	ttcggaacg	gttgagggtc	ccacataatc	agcgcccaca	33540
aaaatcccat	ctcgaacttg	ctcgcgtagg	gagctaaaat	ggccagtata	gccccatggc	33600
accggaacgc	taatctgcaa	gtatatgaga	gccaccccat	tcggcgggat	cacaaaatca	33660
gtcggagaaa	acaacgtata	caccccggac	tgcaaaagct	gttcaggcaa	acgcccctgc	33720
gggtccctctc	ggtaacaccg	caaagcctcg	ggtaaagcag	ccatgccaaag	cgtttaccgt	33780
gccaaagagcg	actcagacga	aaaagtgtac	tgaggcgctc	agagcagcgg	ctatatactc	33840
tacctgtgac	gtcaagaacc	gaaagtcaaa	agttcacccg	gcgcgcccga	aaaaacccgc	33900
gaaaatccac	caaaaaagcc	cgcgaaaaac	acttccgtat	aaaatttccg	ggttaccggc	33960
gcgtcaccgc	cgcgcgacac	gcccgcgccg	ccccgcgctc	ctccccgaaa	cccgcgcgcg	34020
ccacttccgc	gttccaaga	caaaggctcg	gtaactccgc	ccacctcatt	tgcatgttaa	34080
ctcggtcgcc	atcttgcggt	gttatattga	tgatg			34115

<210> 7
<211> 44
<212> DNA
<213> Artificial

<220>
<223> primer P5L

<400> 7
gcgcacgcgt ctctatcgat gaattccatt ggtgatggac atgc 44

<210> 8
<211> 36
<212> DNA
<213> Artificial

<220>
<223> primer P5ITR

<400> 8
gcgcatttaa atcatcatca ataataacc tcaaac 36

<210> 9
<211> 31
<212> DNA
<213> Artificial

<220>
<223> primer P5XTOP

<400> 9
gatacctagg aacgaggagg atttgatatt g 31

<210> 10
<211> 20
<212> DNA
<213> Artificial

<220>
<223> primer P5XBOT

<400> 10
atgtacgcct ccgcgctcac 20

<210> 11
<211> 31
<212> DNA
<213> Artificial

<220>
<223> primer P5E4

<400> 11
gatcgaattc ccactctgta ccccatctct g 31

<210> 12
<211> 31967
<212> DNA
<213> Simian adenovirus

<220>

<221> CDS
<222> (13796)..(15322)

<220>
<221> CDS
<222> (18257)..(21010)

<220>
<221> CDS
<222> (27192)..(29015)

<400> 12
catcatcata atatacctta tttgggaacg gtgccaatat gataatgagg aggcgggggtt 60
aggggtggag tgagggtggg gtgcggatga cgcgggcgcg gggcgggggtg ggagtctgac 120
gtggggcgcg ggggtggagcg cgagggtgag ggcggggcga gggcggcggg cgcggcggaa 180
ttgacgtaca cggtagtaag tttgagcggg aattaagtga attgggcgtg ttttttgtaa 240
ctttttgacg tacacggtag taagtgtgag cggaaattaa gtgaattggg cgtgtttttt 300
gtaacttttt gacgtacacg gtagtaagtt tgagcggaaa ttaagtgaat tgggcgtggt 360
ttttgtaact ttttgacgta cacggtagta agtttgagcg gaaattaagt gaattgggcg 420
tgttttttgt aactttttga cgtacacggt agtaagtttg agcggaaatt aagtgaattg 480
ggcgtgtttt ttgtaacttt ttgacgtaca cggtagtaag tttgagcggg aattaagtga 540
attgggcgtg ttttttgtaa ctttttgacg tacacggtag taagtgtgag cggaaattaa 600
gtgaattggg cgtgtttttt gtaacttttt ggtcattttg gcgcgaaaac tgagtaatga 660
ggaagtgaga cggactctgc ctttttttac ggttgggagg gaaaactgct gatcagcgct 720
gaactttggg ctctgacgcg gtggtttccc tacgtggcag tgccacgaga aggctcaaag 780
tcctcgtttt attgtgtgct cagccttttt gaggggtattt aaacaccgtc agaccgtcaa 840
gaggccactc ttgagtgcga gcgagtagag ttttctcctc cgtcgtctgcc gcggctgctc 900
agtcttaccg ccaggatgcg aatgctgccg gagatcttca ccgggtcctg ggaagatggt 960
ttccagggac ttttagaatc tgaagacaac tttccccaac ctcttgagcc ggaggagcta 1020
cctgaggttt cgcttcacga tctgtttgac gtggagggtg agagccccga cggagatccg 1080
aacgaggaag ctgttgatgg tatgttcccc gactggatga tatctcagag cgagagtgtc 1140
gaaggcagtg cggactcggg cgtttctggg gttggaaacc tgggtggagg ggatctggac 1200
ttgaagtgtt acgaggaagg ttttctcct agcgactcag agactgatga agcctcagaa 1260
gcggaaggtc aagaggagtc tgtgtgtggt tatgtgaaga ttaatgagg ggagaacctg 1320
ctggtgttg actgtccgga ccaacctgga catggctgtc gagcctgtga ctttcaccgg 1380
gggaccagcg gaaacccgga agctatctgt gctttgtgct acatgcgtct gaacgagcac 1440
tgcatataca gtgagtgtta ttcattgggt atttatgggg aaagttgggg gaaagtcttg 1500
agaaggggaa aagtttaaca tgtcattttt gtacttgata ggtccagttt cggacgctga 1560
gggggattct gaggccccctg ctggtccttc ccagccctca ccctgctctt tgaccgccac 1620
gcccgcacct gacctagtta gaccaacgcc ctgccgagtg tcctgtagac gacgtgcagc 1680
tgttaattgc atagaagatt tattggcccc tgatgacgag aacgcacctt tgaacctgtg 1740

cctgaaacgc	cctaagacat	cttgagtgtt	tatgctgtta	ataaaagtgt	tgacccttag	1800
atcctgtgtt	tattccttgg	gcgtgtgctg	gggtatataa	agcagctgctg	ggctggagtg	1860
ttagttttatt	ctgatggagt	actggagtga	gctgcagaat	taccagagcc	tccggcgctt	1920
gctggagttg	gcctctgcca	gaacatccac	ctgctggagg	ttctgttttg	gctcgactct	1980
cagtaacgtg	gtgtatcggg	tgaagcaaga	gtacagctcg	cgcttttctg	agctgttggc	2040
ccgctacccg	gctgtttttg	tttctctgga	tctaggccat	cacgtttatt	tccaagaagc	2100
tgtagtcaga	tatttgatt	tttctactcc	cgggcgtgctg	gtttctgcga	ttgccttcat	2160
ctgctttgtg	ctagatcgat	ggagcgccca	aaccgcctg	agcccggggt	acaccctgga	2220
ctacctgacc	atgtccctgt	ggagggccat	gctgcggaag	aggaggggtct	caggcttctc	2280
gccggcgctg	cctccgcacg	gactggatcc	ggtgctggag	gagtcggagc	tggaggagga	2340
ggagaacccg	agggccggcc	tggaccctcc	ggcggaatag	tgacggaacc	ggaggatccc	2400
caagagggta	ctagtcaggg	gggagggggg	ccgaagagaa	agcgggatga	agaggaggcg	2460
atggaccccg	acaggtttct	aaaagaactg	actttaagct	taatgtctaa	gagaagaccc	2520
gagacggtgt	ggtggtctga	tttgagaaag	gagttccacc	agggggagat	gaatctgttg	2580
tacaagtatg	ggtttgagca	ggtgaagact	cactggctgg	aagcctggga	ggactgggag	2640
atggctttta	acatgtttgc	caaggtggcg	ctgcgcccgg	acactattta	caccgtgact	2700
aagacggtgg	aaatccgcaa	gcctgtgtat	gtgattggca	acggggccgt	ggttcggttc	2760
cagaccaccg	accgggtggc	ctttaattgc	tgtatgcaga	acctgggccc	gggggtgatt	2820
aatcttaatg	gagtgacctt	ttgcaatgtc	agattcgctg	gggatggatt	caacgggacg	2880
gtgtttgccg	ccaccacca	gataacccta	cacggggtgt	tcttcagca	tgtaggcggg	2940
gcttgtgtag	atacctgggc	gagggcctct	gtgaggggct	gcacctttgt	gggctgttgg	3000
aaagcgggtg	tgggtcgacc	caagagtgtg	ctgtctgtga	agaaatgtgt	gtttgagaga	3060
tgtctgatgg	ccatggtggt	ggagggccag	ggtaggatcc	gccataacgc	gggctccgag	3120
aatacctgtt	ttgccctgct	gaagggtagc	gcgaccgtga	agcataacat	gatctgcggg	3180
gtgggtcact	cgcagctgct	gacctgtgctg	gatggcaact	gccaggccct	gcgcacggtg	3240
catgtggtgt	cccaccggcg	ccgcccctgg	ccggtgtttg	aacataacat	gctgatgcgc	3300
tgtaccatgc	acctgggcta	ccgccgcggc	gtgtttgtgc	cccatcagtg	taacctgacc	3360
cacaccaagg	tgttgctgga	gacggatgct	ttttcgcgag	tgaatctgaa	tgggggtgttc	3420
gatctgacta	tggagatgta	caagatagtg	agattttgatg	aatcaaagac	ccgttgtcgc	3480
ccctgcgagt	gcggtgccaa	tcacctgagg	atgtatcccg	tgaccctgaa	cgtgacggag	3540
gagctgcgcc	cggaccacca	gatgctgtcc	tgtctgcgca	ccgattacga	aagcagcgat	3600
gaggattaag	aggtgagggg	cggggcttgc	atgggggtata	aaggtggggg	aggaggtggg	3660
gagggggaaa	acccaaaatg	agcggatcga	tggaaaggag	cgctgtgagt	tttgagggcg	3720
gggtgttcag	cccatatctg	acaacccgtc	tccccgcctg	ggcaggagtg	cgtcagaatg	3780

tggtgggctc	caacgtggac	ggacgtccgg	tggcccctgc	caactccgcc	actctcacct	3840
acgccaccgt	cggatcgtcg	ctggacaccg	ccgctgccgc	cgccgcttca	gccgccgctt	3900
ctactgctcg	cggtatggca	gctgatttcg	gactgtatca	gcaactggct	gcgcctcgct	3960
cgtcgctgag	agaagatgat	gccctgtccg	tggtgctgac	ccgcctggag	gagctgtccc	4020
agcagctgca	agagctgtct	gccaaagtgg	atgcacagaa	cgtccccgct	acccaatgaa	4080
taaataaacg	agacaccgag	tgtgttttga	aatcaaaatg	tgtttttatt	tgttttttct	4140
ggcgcggtag	gcccttgacc	acctgtcgcg	gtcgttaagg	accttgtgga	tgttttccag	4200
cacccggtag	aggtgggctt	ggatgttgag	gtacatgggc	atgagcccgt	ctcgggggtg	4260
gaggtagcac	cactggaggg	cgtcgtgctc	gggggtggtg	ttgtagataa	tccagtcgta	4320
gcagggtttt	tgggcatgga	agcgggaagat	gtctttgaga	agcaggctga	tggccagggg	4380
gaggcccttg	gtgtaggtgt	tcacaaagcg	gttgagctgg	gagggatgca	tgcgggggga	4440
gatgagatgc	atcttggcct	gaatcttgag	gttggcgatg	ttgccgccca	gatcccgccg	4500
ggggctcatg	ttgtgcagga	ccaccaggac	ggtgtagccg	gtgcacttgg	ggaatttgtc	4560
atgcaacttg	gaaggggaagg	cgtggaagaa	cttggagacc	cccttgtggc	cgccgagggt	4620
ctccatgcat	tcgtccatga	tgatggcgat	gggacccttg	gcggccgccc	tggcgaagac	4680
gttgtcgggg	tgggagacgt	cgtagtcttg	ttccagggtg	agctcgtcgt	aggccatttt	4740
gacgaagcgg	gggagcaggg	tgcccgactg	ggggacgatg	gtaccttcgg	gacccggggc	4800
gtagttagccc	tcgcagattt	gcattctcca	ggccttgatc	tccgaggggg	ggatcatgtc	4860
cacctggggc	gcgatgaaga	agacgggtctc	cggggcgggg	ttgatgagct	gggaggagag	4920
gaggttgcg	agcagctgcg	acttgccgca	cccgggtggg	ccgtagatga	ccccgatgac	4980
gggttgacgc	tggtagttaa	aggagctgca	gctgccgtcc	tcgcgcagga	acggggcgac	5040
ctcgttcac	atgcttctga	cgtgatgggt	ttccctgacg	aggtcttgca	agagccgctc	5100
gccgcccagg	gagagaagct	cttcagggt	gcggaaatgc	ttgaggggtt	tgaggccgtc	5160
ggccatggtc	atcttttcca	gggactggcg	gagcaggtag	aggcgggtccc	agagctcggt	5220
gacgtgttct	acggcatctc	gatccagcag	acttcttggt	tgcggggggtt	ggggcggtt	5280
tggctgtagg	ggaccagccg	gtgcgcgtcc	agggaggcga	gggtgacgtc	tttccagggc	5340
cgcagcgttc	gcgtgagggt	ggtctcggtg	acggtgaagg	gatgcgctcc	cggttggggc	5400
ctggccaggg	tcctcttgag	actcatcctg	ctggtgtgga	agcgggcgtc	ttctccctgg	5460
gagtcggcca	ggtagcattt	gagcatgagg	tcgtagctga	ggcctcggc	cgcgtggccc	5520
ttggcgcgca	gcttgccctt	ggagacgtgt	ccgcaggcgg	gacagtgcag	gcacttgagg	5580
gcgtagagct	tgggggcccag	gaagacggac	tcgggggagt	aggcgtcggc	gccgcactga	5640
gcgcacgtgg	tctgcactc	gacgagccag	gtgagctccg	ggtgttgggg	atcaaaaacc	5700
agctggcccc	cgtgtttttt	gatgcgcttc	ttacctcggg	tctccatgag	gcggcgctcc	5760
gcttcggtga	cgaagaggct	gtcgggtgtc	ccgtagacgg	atttgagcgc	gcgctgctcc	5820
aggggaatcc	cgcgatcctc	cgcgtgcagg	aactcggacc	actctgagac	gaaggccccg	5880

gtccacgcga	ggacaaagga	ggcgatctgg	gacgggtagc	ggtcgttctc	caccagggga	5940
tccaccttct	ccaggggtgtg	caggcagagg	tcgtcctcct	ccgcgtccat	gaagggtgatt	6000
ggcttgtaag	tgtatgtcac	gtgaccgtcg	gggtcgcgcg	tgggcttata	aaagggggcg	6060
tgcccggcct	ccccgtcact	ttcttccgca	tcgtctgtga	cgagatccag	ctgctcgggt	6120
gagtaggcgc	gctggaaggc	gggcatgacc	tcggcgctga	gggtgtcagt	ttccacgaac	6180
gaggtggatt	tgatattgac	ctgtccggcg	gcgatgcttt	tgacggtggc	ggggtccatc	6240
tggtcagaaa	agacgatctt	tttgttgctc	agcttggtgg	cgaacgaccc	gtagagggcg	6300
ttggagagca	gcttggcgat	ggagcgcagg	gtctggttct	tctcgcggtc	ggcgcgctcc	6360
ttggcggcga	tgttgagctg	gacgtactcg	cgggccacgc	agcgccattc	ggggaagacg	6420
gtggcgcgct	cgtccggcag	gaggcgcacg	cgccagccgc	ggttggtgag	ggtgatgagg	6480
tccacgctgg	tggccacctc	gccgcgcagt	ggctcgttgg	tccagcagag	gcgcccgcgc	6540
ttgcgcgagc	agaagggggg	caggacgtcg	agctggctct	ccgcgggggg	gtcggcgctc	6600
atggtgaaga	tgcccggtag	caggtggcgg	tcgaagtagt	cgatggcgac	cgcggggtcg	6660
gcgagggcgc	gttcccagtc	cctgaccgcc	agggcgcgct	cgtaggggtt	gaggggcgcc	6720
ccccagggca	tgggatgggt	gagggccgag	gcgtacatgc	cgcatatgtc	gtagacgtag	6780
aggggctcgc	ggagcacgcc	gaggtaggtg	ggatagcagc	gtccgccgcg	gatgctggcg	6840
cgcacgtagt	cgtacatctc	gtgcgagggg	gcgaggaggc	cgctccgag	gtcgccgcgc	6900
tgcggtctga	cggcccggta	ggtgacctgg	cggaagatgg	cgtgcgagtt	ggaggagatg	6960
gtgggccgct	ggaagatggt	gaagctggcc	tcggggagtc	cgacggcgtc	gtggacgaac	7020
tgggcgtagg	agtcgcgcag	cttctgcacg	agcgcggcgg	tgacgagcac	gtccagggcg	7080
cagtagtcga	gggtctcgcg	gacgaggtcg	taacggggct	cttgcttctt	ttcccagagt	7140
tcgcggttga	ggaggtactc	ctcgcgatcc	ttccagtact	cttcggccgg	aaagccgcgt	7200
tcgtccgcca	ggtaagaacc	cagcatgtag	aagcggttga	cggctcggta	gggacagcag	7260
cccttctcga	cgggcagggg	gtaggcctgc	gcggccttcc	tgagcgaggt	gtgggtgagg	7320
gcgaagggtg	cgcgaccat	gaccttgagg	aactggaacc	tgaagtcggt	gtcgtcgcag	7380
gcgccccgct	cccagagccc	gtagtcggtg	cgtttctggc	tgcgggggtt	gggcagggcg	7440
aagggtgacgt	cgttgaagag	gatcttgccg	gcgcgcggca	tgaagtgcg	ggtgatcctg	7500
aagggccccg	gcacgtccga	gcggttggtt	atgacctggg	ccgcgaggac	gatctcgtcg	7560
aagccgttga	tgttggtggc	gacgatgtag	agctcgacga	agcgcgggcg	cccctgcagc	7620
ttgggggcct	tcttgagctc	ctcgtaggtg	aggcagtcgg	gcgagtagag	gcccagctcc	7680
tgtcggggcc	attcgccac	ctgggggttg	gcttgcaaga	agccccgcca	gagctgcagg	7740
gcgagctggg	tctggaggcg	gtcgcggtag	tcgcggaaact	ttttgcccac	cgccatcttc	7800
tcgggggtga	ccacgtagaa	ggtgcggccg	tcctggcccc	aggcgtccca	gttctgctcg	7860
cgggcgagac	ggcaggcctc	ctcgacgagg	gcctcctccc	cggagagatg	catgactagc	7920

atgaagggga	cgagttgctt	gccgaaggca	cccatccacg	tgtaggtctc	tacgtcgtag	7980
gtgacgaaga	gacgttcggt	gcgaggatgc	gagccgagag	gaaagaagtt	gatctcctgc	8040
caccagcccg	aggagtgggc	gttgacgtgg	tggaagtaga	agtcacgccg	gcggaccgtg	8100
cattcgtgct	gatatttgta	aaagcgggcg	cagtactcgc	agcgctgcac	gctctgcact	8160
tcctgaacga	gatgcacccg	gcgcccgcgc	accaggaggc	ggagggggca	gtccagtggg	8220
gcttcggcgc	gctgtccttc	agcctcgtca	tgtctttctg	cacctgcacg	ctcctgctgt	8280
gggtggagga	cggagggagt	gacgacgccg	cgcgagccgc	aggtccagat	gtcgacgcgc	8340
ggcggcctga	ggctcagcgc	caggggtgcg	atctgagcgg	cgtccaggga	gtcgaggaag	8400
gcctcgtga	ggctcagcgg	cagcgtccgc	cggtggactt	gcaggagacg	ggtaagggcc	8460
ggcgccaggc	gctgatggta	cttgatctcg	agcggttcgt	tggtaggagt	gtcgatggcg	8520
tagagcaggg	cctgaccgcg	ggcggcgacg	atggtgccgc	ggtgccggcg	gtaggtggcg	8580
tattcggggg	ggctcggttac	atcacccgcc	tgggcctggc	gccgggcggc	agcggggggt	8640
ctggtcccgc	cggcatgggc	ggcagcggca	cgtcggcgcg	gggctccggc	agcggctggt	8700
gctgagctcg	cagctgactg	gcgtgcgcga	cgacgcggcg	gttgaggtcc	tggatgtgcc	8760
tccgctgcgt	gaagaccacc	ggtccccgga	ctcggaacct	gaaagagagt	tcgacagaat	8820
caatctcggc	atcgttgacg	gccgcctgac	gcaggatctc	ctgcacgtcg	cccagattgt	8880
cctggtaggc	gatctcggac	atgaactggg	cgatctcttc	ctcctggagt	tcgccgcgtc	8940
cggcgcgctt	gacggtggcc	gcgaggtcgt	tggagatgcg	agccatgagc	tgggagaagg	9000
cgttgaggcc	gttctcgttc	cacacgcgac	tgtagacgac	gttgccgacg	gcgtcccggg	9060
cgcgcacgac	cacctgcgcg	acgttgagct	ccacgtgtcg	cgcaagacg	gcgtagtgtc	9120
gcaggcgctg	gaagaggtag	ttgaggggtg	tggcgaatgt	ctcgacagac	aagaagtaca	9180
tgaccacagc	gcgcagcgtc	atctcgttga	tgtctccgag	ggcttccaag	cgctccatgg	9240
cctcgtagaa	gtcgacggcg	aagttgaaga	actgggagtt	gcgcgccgcg	accgtcagct	9300
cgtcttgcaa	gagccggatc	agctgggcca	cggctctccc	cacctcgcgt	tcgaaggccc	9360
ccggcgcttc	ttcctcctct	ggttcctcgg	cggcctcttc	ttccatgacg	gcttcctctt	9420
cctccggttc	ctcgggcacg	ggcctccggc	ggcgacggcg	cctgatgggc	aggcgggtcca	9480
cgaagcgctt	gatgatctct	ccgcggcgcc	ggcgcatggt	ttcggtgacg	gcgcggccgt	9540
tctctcgggg	ccgcagttcg	aagacgcccc	cgcgcaggcc	gccggcgccg	ccgagagggg	9600
gcaggaggtg	ggggccttcg	ggcagcgaga	gggcgctgac	gatgcaccgt	atcatctgtt	9660
gcgtaggtac	agctctccag	gagtcgttga	gcgagtcacg	ttggacggga	tccgagaact	9720
tttcgaggaa	agcttcgatc	caatcgcagt	cgcaaggtaa	gctgaggacg	gtgggatgag	9780
gggcttgggc	ggaggcgga	gcggcagaag	aggaggagga	gggcaggctg	gaggtgatgc	9840
tgctgatgat	gtaattgaag	taggcggttt	tcaaacggcg	gatggtggcg	aggaggacga	9900
cgtctttggg	cccggcctgc	tggatgcgca	ggcggctggc	catgccccag	gcgtggctct	9960
ggcatcggcg	caggtccttg	tagtagtctt	gcatgagtct	ctcgacgggg	acgtcgtctt	10020

cgtcggcccg gtcggccatg cgggtggagc cgaacccgcg caggggctgc agcagggcca 10080
ggtcggcgac cacgcgttcg gccagcacgg cctgctggat ctgggtgagc gtggtctgga 10140
agtcgtccag gtccacgaag cgggtggtagg agcccgtgtt gatggtgtag gtgcagttgg 10200
ccatgacgga ccagttgacg acttgcatgc cgggctgggt gatctcgggtg tagcggaggc 10260
gcgagtaggc ccgcgactcg aagacgtagt cgttgacaggt gcgcacgagg tactggtagc 10320
cgacgaggaa gtgcggcggc ggctcgcggt agaggggcca gcgcacgggtg gcgggggcg 10380
cgggggcccag gtcctccagc atgaggcggt ggtagtggtg gacgtagcgc gagagccagg 10440
tgatgccggc ggcgagggtg gcggcgcggt cgaagtcgcg gacgcggttc cagatgttgc 10500
gcaagggggc gaagcgctcc atggtgggca cgctctggcc ggtgaggcgg gcgcagtcct 10560
gcacgctcta gacgggacag agagcgggag gttagcggct ccgctccgtg gcctggggga 10620
cagaccgcca ggggtgcgacg gcggggaacc ccggttcgag accggctgga tccgtccgtc 10680
cccgacgcgc cggccccgcg tccacgacct caccagaggc cgagaccag ccgcggtgcc 10740
cggaccccag atacggaggg gagccttttt gtggtttttt cccgtagatg catccggtgt 10800
tgcgacagat gcgtccgtcg ccagcgccgc cgacgcagcc gccgctccc cccccacta 10860
gcgccgcgga ggctctgtcc ggcgcccgcg gcgacccgga ggaggaggcc atcctcgact 10920
tggaagaagg cgagggcctg gcccggctgg gagcgccctc ccccgagcgc catccccgcg 10980
tgcagctggc gagagactcg cgccaggcct acgtgccgcc gcagaatctg ttcagggacc 11040
gcagcggcca ggagcccag gagatgaggg accgcaggtt tcacgcgggg cgggagctgc 11100
gcgcgggctt cgaccgtcg cggtgttgc gcgccgaaga cttcgagccc gacgagcgca 11160
gcggagtaag tccggcacgg gcgcacgtgt cggcggccaa cctggtgacc gcgtacgagc 11220
agacggtgaa cgaggagcgg agctttcaga aaagcttcaa caaccacgtg cgcaccctga 11280
tcgcgcgca ggagggtggc atcggcctga tgcattctgt ggactttgtg gaggcgtagc 11340
tgcagaacct gtcgagcaag ccgctgacgg gcagattgtt cctgatcgtg cagcacagtc 11400
gggacaacga gacgttccgc gaggcgatgc tgaacatcgc ggagcccag ggccgctggc 11460
tcttgacat gattaacatc ctgcagagca tcgtggtgca ggagcgcagc ctgagcctgg 11520
ccgacaaggt ggcggccatc aactacagca tggtgagcct gggcaagttt tacgcccga 11580
agatctacaa gagcccctac gtgcccatag acaaggagggt gaagatcgac agcttttaca 11640
tgcggatggc gctgaaagtg ctgacgctga gcgacgatct gggggtgtac cgcaacgacc 11700
gcatccacaa ggccgtgagc gccagccgcc ggcgcgagct gagcgaccgc gagctgatgc 11760
acagcctgcg gagggcgctg gcgggcgccc gcggcgcgga ggaggccgag tcctacttcg 11820
acatgggggc ggacttgacg tggcagccca gcgcgcgggc cctggaggcg gcgggctacc 11880
gcggcgcgcg cggcgtggtc gaggcggagg acgaggacga ggtggagtac gagggaggag 11940
actgatcggc gaggtgtttt cgtagatgca gcgcgcgacg gcggcgcgga gcgggcccga 12000
gggggacccc gccgtgctgg cggccctgca gagccaacct tcgggcgtga acgcctccga 12060

tgactggg	cg	gccc	at	gg	cc	at	tt	gg	cc	tt	ga	cc	ac	cc	cg	ag	gc	ct	tt	12120									
tagacag	cag	cc	gc	ag	gcc	a	acc	gc	tt	tt	tc	gg	cc	at	ct	tt	ga	ag	cc	gc	tg	tg	tg	gc	cc	tc	gc	g	12180
caccaac	cccc	ac	gc	ac	gaga	ag	gt	c	ct	gg	gc	gg	t	gg	t	ga	ac	gc	gc	tg	tg	tg	tg	tg	tg	tg	tg	tg	12240
gatccg	caag	gac	gag	gc	gg	gc	gt	at	tt	ta	ca	ac	gc	cc	tg	ct	gg	ag	cg	gg	gg	gg	gg	gg	gg	gg	gg	gg	12300
caacag	cacc	aac	gt	gc	agg	cca	ac	ct	gg	ga	cc	gt	ct	ga	cg	ac	gg	ag	cg	tg	gc	gg	gg	gg	gg	gg	gg	gg	12360
ggcgcag	cg	gag	cg	ct	tca	tg	cg	cg	ac	ac	ga	ac	ct	gg	gc	tc	gc	ag	gt	gg	cc	ct	ga	ac	gc				12420
cttcct	gag	c	gc	ag	cc	gg	cca	ac	gt	gg	gc	gc	gg	gg	cg	ag	gag	g	act	ac	gc	ta	g	tt	t	cat			12480
cagcgc	gct	g	cg	ct	c	ct	gg	tg	gg	cc	gag	gt	gc	cg	ag	ag	g	g	g	t	ac	ag	t	cg	gg	t	cc		12540
ggactac	ttc	tt	cc	ag	ac	ct	cg	gc	gg	cg	agg	gc	ct	gc	ag	ag	g	t	ga	ac	ct	ga	cg	ag	gc	ct	t		12600
caagaac	ctg	ga	ag	gc	at	gt	gg	gg	gc	gt	gc	gg	cc	cc	cg	tg	gg	cg	ac	cc	gg	gg	cg	ac	gc	at	ct	c	12660
cagcttg	ctg	ac	gcc	ga	aca	cg	gg	ct	g	ct	g	ct	g	ct	g	ct	at	cg	cg	cc	ct	ta	cca	at	ag				12720
cagtacc	atc	ag	cc	gc	g	act	cg	t	ac	ct	gg	gc	cc	ct	gat	c	ac	gc	t	g	t	ac	gc	ag	gc	cat			12780
cgggcag	gc	cag	gt	gg	ac	g	ag	ca	gc	ct	ta	cc	ag	ga	g	att	ac	gc	ag	cg	tg	gc	gg	gc	cc	ct			12840
ggggcag	cag	gac	ac	gg	g	ta	gc	ct	gg	agg	gc	gac	gc	t	ga	at	tt	ct	g	ct	ga	cca	ac	cg	gc				12900
gcagaag	atc	cc	ct	ccc	ag	t	ac	ac	gc	t	gag	ca	cg	ga	gg	ag	gag	cg	ca	t	ct	tg	cg	t	ac	gt			12960
gcagcag	tcc	gt	gag	cc	tg	t	at	ct	gat	gc	cg	ag	gg	gg	gg	gc	ag	cc	cc	ct	cg	t	cg	gc	t	gga			13020
catgac	ggcc	cg	ta	ac	at	gg	ag	cc	gc	t	gc	t	gc	t	gc	ta	cc	cg	gc	cc	gt	tc	gt	ga	acc	g			13080
cctgat	ggac	ta	ct	gc	ac	gc	gc	gc	gc	gc	cat	ga	ac	gg	gc	gag	t	act	tt	ta	ga	ac	gc	cc	at				13140
cctgaac	ccg	ca	ct	gg	at	gc	gc	cg	t	cc	gg	tt	ct	ac	ac	gc	gg	g	act	tt	tg	ac	at	g	ccc	ga			13200
ggg	gc	ac	gc	ac	gg	gt	tc	ct	gt	gg	ac	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	13260
ggg	ca	ga	ag	gag	gg	gc	gg	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	13320
ccagag	ccc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	13380
gcgccc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	13440
gaacttt	ccc	a	a	ca	ac	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	13500
cgcccag	gag	ca	gc	gg	gg	gag	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	13560
caggtgg	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	13620
cttggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	ggg	13680
gctggg	tc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	13740
gccatg	ggt	cg	ca	gc	gt	cg	gt	cc	ct	tt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	gt	13798
agg	cg	gc	gt	gga	gt	cc	cc	gt	at	gc	tac	gcc	gag	ggt	cct														13846
Arg	Arg	Ala	Val	Gly	Val	Pro	Pro	Val	Met	Ala	Tyr	Ala	Glu	Gly	Pro														
			5					10					15																
cct	cct	tct	tac	gaa	acg	gtg	atg	ggc	gcc	gcg	gat	tcg	ccg	gcc	acg														13894
Pro	Pro	Ser	Tyr	Glu	Thr	Val	Met	Gly	Ala	Ala	Asp	Ser	Pro	Ala	Thr														
			20				25						30																
ctg	gag	gcg	ctc	tac	gtc	cct	ccc	cg	tac	ctg	ggg	cct	acg	gag	ggg														13942

Leu	Glu	Ala	Leu	Tyr	Val	Pro	Pro	Arg	Tyr	Leu	Gly	Pro	Thr	Glu	Gly	
35						40					45					
agg	aac	agc	atc	cgt	tac	tca	gag	ctg	gcg	ccg	ctg	tac	gac	acc	acc	13990
Arg	Asn	Ser	Ile	Arg	Tyr	Ser	Glu	Leu	Ala	Pro	Leu	Tyr	Asp	Thr	Thr	
50				55					60							
cgc	gtg	tac	ctg	gtg	gat	aac	aag	tcg	gcg	gac	atc	gcg	tcg	ctg	aac	14038
Arg	Val	Tyr	Leu	Val	Asp	Asn	Lys	Ser	Ala	Asp	Ile	Ala	Ser	Leu	Asn	
			70					75						80		
tac	cag	aac	gac	cat	agc	aac	ttt	ctg	acc	acg	gtg	gtg	cag	aac	aat	14086
Tyr	Gln	Asn	Asp	His	Ser	Asn	Phe	Leu	Thr	Thr	Val	Val	Gln	Asn	Asn	
			85					90					95			
gac	ttt	acc	ccg	gtg	gag	gcg	ggc	acg	cag	acc	ata	aat	ttc	gac	gag	14134
Asp	Phe	Thr	Pro	Val	Glu	Ala	Gly	Thr	Gln	Thr	Ile	Asn	Phe	Asp	Glu	
		100					105					110				
cgc	tcg	cgg	tgg	ggc	ggc	gac	ctg	aaa	acc	atc	ctg	cgc	acc	aac	atg	14182
Arg	Ser	Arg	Trp	Gly	Gly	Asp	Leu	Lys	Thr	Ile	Leu	Arg	Thr	Asn	Met	
	115					120					125					
ccc	aac	atc	aac	gag	ttc	atg	tcc	acc	aac	aag	ttc	agg	gcc	cgg	ttg	14230
Pro	Asn	Ile	Asn	Glu	Phe	Met	Ser	Thr	Asn	Lys	Phe	Arg	Ala	Arg	Leu	
130				135						140					145	
atg	gta	gag	aaa	gtg	aac	aag	gaa	acc	aat	gcc	cct	cga	tac	gag	tgg	14278
Met	Val	Glu	Lys	Val	Asn	Lys	Glu	Thr	Asn	Ala	Pro	Arg	Tyr	Glu	Trp	
				150					155					160		
ttt	gag	ttc	acc	ctg	ccc	gag	ggc	aac	tac	tcg	gag	acc	atg	acc	ata	14326
Phe	Glu	Phe	Thr	Leu	Pro	Glu	Gly	Asn	Tyr	Ser	Glu	Thr	Met	Thr	Ile	
			165					170					175			
gac	ctg	atg	aat	aac	gcg	atc	gtg	gac	aac	tac	ttg	gaa	gtg	ggg	cgg	14374
Asp	Leu	Met	Asn	Asn	Ala	Ile	Val	Asp	Asn	Tyr	Leu	Glu	Val	Gly	Arg	
		180					185					190				
cag	aac	ggg	gtg	ctg	gag	agc	gac	atc	ggg	gtg	aag	ttt	gac	acg	cgc	14422
Gln	Asn	Gly	Val	Leu	Glu	Ser	Asp	Ile	Gly	Val	Lys	Phe	Asp	Thr	Arg	
	195					200					205					
aac	ttc	cgg	ctg	ggc	tgg	gac	ccg	gtc	acc	aag	ctg	gtc	atg	ccc	ggc	14470
Asn	Phe	Arg	Leu	Gly	Trp	Asp	Pro	Val	Thr	Lys	Leu	Val	Met	Pro	Gly	
210				215						220					225	
gtg	tac	acc	aac	gag	gcc	ttc	cac	ccc	gac	atc	gtc	ctg	ctg	ccc	ggc	14518
Val	Tyr	Thr	Asn	Glu	Ala	Phe	His	Pro	Asp	Ile	Val	Leu	Leu	Pro	Gly	
				230					235					240		
tgc	ggc	gtg	gac	ttc	acg	cag	agc	cgg	ctg	agc	aac	ctg	ctg	ggg	atc	14566
Cys	Gly	Val	Asp	Phe	Thr	Gln	Ser	Arg	Leu	Ser	Asn	Leu	Leu	Gly	Ile	
			245					250					255			
cgc	aag	cgg	atg	ccc	ttc	cag	gcg	ggc	ttt	cag	atc	atg	tac	gag	gac	14614
Arg	Lys	Arg	Met	Pro	Phe	Gln	Ala	Gly	Phe	Gln	Ile	Met	Tyr	Glu	Asp	
		260					265					270				
ctg	gag	ggc	ggc	aac	atc	ccc	gcc	ttg	cta	gac	gtg	gcg	aaa	tac	gag	14662
Leu	Glu	Gly	Gly	Asn	Ile	Pro	Ala	Leu	Leu	Asp	Val	Ala	Lys	Tyr	Glu	
	275					280					285					
gcc	agc	att	cag	aag	gcg	cgg	gag	cag	ggc	cag	gag	atc	cgc	ggc	gac	14710
Ala	Ser	Ile	Gln	Lys	Ala	Arg	Glu	Gln	Gly	Gln	Glu	Ile	Arg	Gly	Asp	
290					295				300					305		
aac	ttt	acc	gtc	atc	ccc	cgg	gac	gtg	gag	atc	gtg	ccc	gtg	gag	aag	14758
Asn	Phe	Thr	Val	Ile	Pro	Arg	Asp	Val	Glu	Ile	Val	Pro	Val	Glu	Lys	

310										315					320					
gat	agc	aag	gac	cgc	agt	tac	aac	cta	ctc	ccc	ggc	gac	cag	acc	aac	14806				
Asp	Ser	Lys	Asp	Arg	Ser	Tyr	Asn	Leu	Leu	Pro	Gly	Asp	Gln	Thr	Asn					
			325					330					335							
acg	gcc	tac	cgc	agc	tgg	ttc	ctg	gcc	tac	aac	tac	ggc	gac	ccc	gag	14854				
Thr	Ala	Tyr	Arg	Ser	Trp	Phe	Leu	Ala	Tyr	Asn	Tyr	Gly	Asp	Pro	Glu					
		340					345					350								
aag	ggc	gtc	agg	tcc	tgg	acg	ctg	ctg	acc	acc	acg	gac	gtc	acc	tgc	14902				
Lys	Gly	Val	Arg	Ser	Trp	Thr	Leu	Leu	Thr	Thr	Thr	Asp	Val	Thr	Cys					
	355					360					365									
ggc	tcg	cag	cag	gtg	tac	tgg	tcg	ctc	ccg	gac	atg	atg	caa	gac	ccc	14950				
Gly	Ser	Gln	Gln	Val	Tyr	Trp	Ser	Leu	Pro	Asp	Met	Met	Gln	Asp	Pro					
370					375					380					385					
gtg	acc	ttc	cgg	ccc	tcc	agc	caa	gtc	agc	aac	tac	ccc	gtg	gtg	gga	14998				
Val	Thr	Phe	Arg	Pro	Ser	Ser	Gln	Val	Ser	Asn	Tyr	Pro	Val	Val	Gly					
				390					395					400						
gtc	gag	ctc	ctg	ccg	gtg	cac	gcc	aag	agc	ttt	tac	aac	gag	cag	gcc	15046				
Val	Glu	Leu	Leu	Pro	Val	His	Ala	Lys	Ser	Phe	Tyr	Asn	Glu	Gln	Ala					
			405					410					415							
gtc	tac	tcg	cag	ctc	atc	cgc	cag	tcc	acc	gcg	ctc	acg	cac	gtc	ttc	15094				
Val	Tyr	Ser	Gln	Leu	Ile	Arg	Gln	Ser	Thr	Ala	Leu	Thr	His	Val	Phe					
		420					425					430								
aac	cgc	ttc	ccc	gag	aac	cag	atc	ctg	gtg	cgc	ccg	ccc	gct	ccg	acc	15142				
Asn	Arg	Phe	Pro	Glu	Asn	Gln	Ile	Leu	Val	Arg	Pro	Pro	Ala	Pro	Thr					
	435					440					445									
att	acc	acc	gtc	agt	gaa	aac	gtt	ccc	gcc	ctc	aca	gat	cac	gga	acc	15190				
Ile	Thr	Thr	Val	Ser	Glu	Asn	Val	Pro	Ala	Leu	Thr	Asp	His	Gly	Thr					
450					455					460					465					
ctg	ccg	ctg	cgc	agc	agt	atc	agt	gga	gtc	cag	cgc	gtg	acc	atc	act	15238				
Leu	Pro	Leu	Arg	Ser	Ser	Ile	Ser	Gly	Val	Gln	Arg	Val	Thr	Ile	Thr					
				470				475						480						
gac	gcc	cgg	cga	agg	acc	tgc	ccc	tac	gtg	cac	aag	gcc	ctg	ggc	ata	15286				
Asp	Ala	Arg	Arg	Arg	Thr	Cys	Pro	Tyr	Val	His	Lys	Ala	Leu	Gly	Ile					
			485					490					495							
gtc	gct	ccc	aaa	gtg	ctc	tct	agc	cgc	acc	ttt	taa	caagcatg	g			15332				
Val	Ala	Pro	Lys	Val	Leu	Ser	Ser	Arg	Thr	Phe										
		500					505													
cattctcatc	tcgcccagaca	acaacaccgg	ctggggcctg	cgctcggccg	gcatgtacgg											15392				
cggcgcgaag	cggcgctcca	gcgagcacc	cgtccgcgtc	cgcgccact	accgggcccc											15452				
ctggggcgcc	cacaagcgcg	gcgtctccac	gcgcaccacc	gtcgacgacg	ccatcgacgc											15512				
cgtcgtggcc	caggccagac	gctaccgccg	gcccagtcg	acggtggacg	ccgtcatcga											15572				
cagcgtggtg	gccgacgcgc	ggcgatacgc	tcgacgcaag	cgcgctctgc	accgccgtcg											15632				
ccgtcccacc	gccgccatgc	tggccgccag	agcggtcctg	agacgcgcgc	gccgcgtggg											15692				
ccgccgagcc	atgcgccgag	ccgcggccaa	cgccagcgcg	ggcgcgccc	gtcgtcaggc											15752				
cgcccggcag	gccgccgccg	ccatcgccaa	cctggcccaa	ccccgccggg	gaaacgtgta											15812				
ctgggtg	gacgcgtcgg	gcgtgcgcgt	gccggtgcgc	acccgcccc	ctcggagtta											15872				
gaagacaaaa	agacggacga	agactgagtt	tccctgtcgt	tgccagcatg	agcaagcgca											15932				

agttcaaaga agagctgctg gaggccctcg tgcccagagat ctacggccccg gccgccgctg 15992
ccgccgcggt ggcggacgtc aagcccgaag ttaagccccg cgcgctgaag cgggttaaaa 16052
agcgggaaaa gaaagaggag aagcaggaag cagggttgct agacgtcgac gacggcgtgg 16112
agtctgtgcg gtccttcgcg ccccgctcgcc ggggtgcagtg gcgggggtcgc cgcgtcaagc 16172
tcgtccccgcg gccgggcacc gtggtgtctt tcacccccgg cctgcgttcg gccacgcgcg 16232
gcctgaagcg cgagtacgac gaggtctatg gcgacgaaga catcctggag caggccgccc 16292
agcagctcgg ggagtttgct tacggcaagc gcggccgcta cggggagggtg gcgctggcgc 16352
tggaccaggg caatcccacg cccagcctca agcccgtcac gctgcagcag gtgctgccccg 16412
tgagcgcgtc gaccgagagc aagcggggca tcaagagggga gatgggagc ctgcagccca 16472
ccatgcaact catggtgccc aaacggcaga agctggagga cgtgctggag aacatgaaag 16532
tggatcccag catcgagccc gaagtgaag tgcgacccat caaggaagtg ggcccgggccc 16592
taggcgtgca gacggtggac attcagatcc ccgtgcgcgc ctccccggt tctgccacca 16652
ctacgacggc cgtggaggcc atggaaacgc agacggagct gcccgcgcc ttggcggcag 16712
ccgccaccgc cgccgcggct acccgagaga tgggcatgca gaccgacccc tgggtacgagt 16772
tcgccggccc cgccgctcgt ccacgagccc gtcggtacgc ggcgaccacc tcccggctcc 16832
ctgactacgt cttgcatcct tccatcacgc cgacgccccg ctaccgcgga acgaccttc 16892
gccccggtcg cgcgcgcacc accacccgcc gtcgtcgcac caccgcccgc cgtcgcagcc 16952
gtcgcgcact ggctcccatc gcggttcgcc gcgtcgtccg ccgggggtcgc acgctgaccc 17012
tgcccaccgc gcgttaccac cccagcatcg tcatttaacc tgcgctgccg ttttgagat 17072
ggctctgacg tgccgctttc gcttccccgt tcggcactac cgaggaagat ctcgccgtag 17132
gactggtcta gcgggcagcg gtctccgacg ccgccgccgc gcggtgcacc ggcgcagtga 17192
gggcgccatt ctgcccgcgc tgatccccat tatcgccgcc gccatcgggg cgatccccgg 17252
cgtggcctcg gtggccttgc aagcagctcg caaaaattaa ataaagaagg cttgacactc 17312
actgcctggc cctgactgtt tcatgcagac aagacatgga agacatcaat tttgcgtcgt 17372
tggccccgcg gcacggctcg cggccgttca tgggcacctg gaacgagatc ggcaccagcc 17432
agctcaacgg gggcgctttc agttggagca gcctgtggag cggcattaaa aactttgggt 17492
ccacgattaa gacctatggc aacaaggcgt ggaacagtag cactggctcag atgctccgcg 17552
ataagctgaa ggaccagaac ttccagcaga aagtggtaga cggctctggcc tcgggcatca 17612
acgggggtggg ggacctggcc aaccaggcgg tgcagaacca gatcaaccag cgtctggaga 17672
acagccgcca gccgcccgcg gccctgcagc agcgtccgca ggtggaggag gtggaagtgg 17732
aggagaagct gccgcccctg gagacgggtg cgccgggtggg cgtgcctagc aagggggaga 17792
agcggccgcg gcccgagctc gaggagaccc tagtgaccga gaccctggag ccgccctcgt 17852
acgagcaggc cttgaaagag gggggccacgc ccctgccccat gaccgggccc atcggaacca 17912
tggccccgacc ggtctacggc aaggaacaca aagccgtgac gctagagctg cctccgccgg 17972

Page 114

Asn	Thr	Val	Asp	Leu	Gln	Phe	Phe	Ala	Ser	Ser	Ser	Ala	Thr	Ser	Thr	
			740					745					750			
cct	aaa	gcc	gta	ctc	tat	gcc	gag	gac	gtg	gca	ata	gaa	gca	cca	gac	19036
Pro	Lys	Ala	Val	Leu	Tyr	Ala	Glu	Asp	Val	Ala	Ile	Glu	Ala	Pro	Asp	
		755					760					765				
acc	cat	ttg	gtg	tac	aaa	ccg	gca	gtt	aca	acc	acg	acc	act	agt	tcc	19084
Thr	His	Leu	Val	Tyr	Lys	Pro	Ala	Val	Thr	Thr	Thr	Thr	Thr	Ser	Ser	
	770					775					780					
caa	gac	ctg	cta	act	cag	cag	gct	gct	ccc	aac	cga	ccc	aac	tac	att	19132
Gln	Asp	Leu	Leu	Thr	Gln	Gln	Ala	Ala	Pro	Asn	Arg	Pro	Asn	Tyr	Ile	
785					790					795					800	
ggc	ttc	agg	gat	aat	ttt	atc	ggg	ctc	atg	tat	tac	aac	tcc	act	ggc	19180
Gly	Phe	Arg	Asp	Asn	Phe	Ile	Gly	Leu	Met	Tyr	Tyr	Asn	Ser	Thr	Gly	
				805					810					815		
aat	atg	ggg	gtt	ttg	gca	ggg	caa	gct	tct	cag	cta	aac	gcg	gtg	gtt	19228
Asn	Met	Gly	Val	Leu	Ala	Gly	Gln	Ala	Ser	Gln	Leu	Asn	Ala	Val	Val	
			820					825					830			
gac	ttg	caa	gac	aga	aac	acc	gag	ctg	tcc	tac	cag	ctc	atg	ctt	gat	19276
Asp	Leu	Gln	Asp	Arg	Asn	Thr	Glu	Leu	Ser	Tyr	Gln	Leu	Met	Leu	Asp	
		835					840					845				
gct	ttg	ggc	gac	cgc	agt	cgt	tac	ttc	tcc	atg	tgg	aac	cag	gcc	gta	19324
Ala	Leu	Gly	Asp	Arg	Ser	Arg	Tyr	Phe	Ser	Met	Trp	Asn	Gln	Ala	Val	
	850					855					860					
gac	agc	tat	gac	cct	gat	gtc	aga	att	att	gaa	aat	cat	ggg	gtg	gag	19372
Asp	Ser	Tyr	Asp	Pro	Asp	Val	Arg	Ile	Ile	Glu	Asn	His	Gly	Val	Glu	
865					870					875					880	
gat	gag	ctg	cca	aac	tac	tgt	ttc	ccg	cta	gga	ggg	tcg	cta	gta	act	19420
Asp	Glu	Leu	Pro	Asn	Tyr	Cys	Phe	Pro	Leu	Gly	Gly	Ser	Leu	Val	Thr	
				885					890					895		
gaa	act	tat	aca	ggc	cta	tca	ccc	caa	aac	gga	agt	aac	acg	tgg	aca	19468
Glu	Thr	Tyr	Thr	Gly	Leu	Ser	Pro	Gln	Asn	Gly	Ser	Asn	Thr	Trp	Thr	
			900					905					910			
acc	gac	agc	acc	acc	tat	gca	act	aga	ggg	gtg	gaa	atc	ggc	tct	ggc	19516
Thr	Asp	Ser	Thr	Thr	Tyr	Ala	Thr	Arg	Gly	Val	Glu	Ile	Gly	Ser	Gly	
		915					920					925				
aac	atg	ttc	gcc	atg	gaa	att	aat	ttg	gcg	gcc	aat	cta	tgg	agg	agt	19564
Asn	Met	Phe	Ala	Met	Glu	Ile	Asn	Leu	Ala	Ala	Asn	Leu	Trp	Arg	Ser	
	930					935					940					
ttc	ctg	tac	tcc	aac	gtg	gcc	ctg	tac	ctg	ccc	gac	gag	tac	aag	ctc	19612
Phe	Leu	Tyr	Ser	Asn	Val	Ala	Leu	Tyr	Leu	Pro	Asp	Glu	Tyr	Lys	Leu	
945					950					955					960	
acc	ccc	gac	aac	atc	acc	ctc	ccc	gac	aac	aaa	aac	act	tac	gac	tac	19660
Thr	Pro	Asp	Asn	Ile	Thr	Leu	Pro	Asp	Asn	Lys	Asn	Thr	Tyr	Asp	Tyr	
				965					970					975		
atg	aac	ggc	cgc	gtg	gcc	gcc	ccc	agc	tcc	ctc	gac	acc	tac	gtc	aac	19708
Met	Asn	Gly	Arg	Val	Ala	Ala	Pro	Ser	Ser	Leu	Asp	Thr	Tyr	Val	Asn	
			980					985					990			
atc	ggg	gcg	cgc	tgg	tcc	ccc	gac	ccc	atg	gac	aac	gtc	aac	ccc	ttc	19756
Ile	Gly	Ala	Arg	Trp	Ser	Pro	Asp	Pro	Met	Asp	Asn	Val	Asn	Pro	Phe	
		995					1000					1005				
aac	cac	cac	cgc	aac	gcg	gga	ctg	cgc	tac	cgc	tcc	atg	ctg	ctg		19801
Asn	His	His	Arg	Asn	Ala	Gly	Leu	Arg	Tyr	Arg	Ser	Met	Leu	Leu		

1010						1015						1020						
ggc Gly	aac Asn 1025	ggc Gly	cgc Arg	tac Tyr	gta Val	ccc Pro 1030	ttc Phe	cac His	atc Ile	caa Gln	gtg Val 1035	ccc Pro	cag Gln	aaa Lys	19846			
ttc Phe	ttc Phe 1040	gcc Ala	atc Ile	aaa Lys	aac Asn	ctc Leu 1045	ctg Leu	ctc Leu	ctc Leu	ccc Pro	ggg Gly 1050	tcc Ser	tac Tyr	acc Thr	19891			
tac Tyr	gag Glu 1055	tgg Trp	aac Asn	ttc Phe	cgc Arg	aag Lys 1060	gac Asp	gtc Val	aac Asn	atg Met	atc Ile 1065	ctc Leu	cag Gln	agc Ser	19936			
agc Ser	ctg Leu 1070	ggt Gly	aac Asn	gac Asp	ctc Leu	cgc Arg 1075	gtc Val	gac Asp	ggg Gly	gcc Ala	agc Ser 1080	gtc Val	agg Arg	ttc Phe	19981			
gac Asp	agc Ser 1085	atc Ile	aac Asn	ctg Leu	tac Tyr	gcc Ala 1090	aac Asn	ttc Phe	ttc Phe	ccc Pro	atg Met 1095	gcc Ala	cac His	aac Asn	20026			
acc Thr	gcc Ala 1100	tcc Ser	acg Thr	ctc Leu	gag Glu	gcc Ala 1105	atg Met	ctg Leu	cgc Arg	aac Asn	gac Asp 1110	acc Thr	aac Asn	gac Asp	20071			
cag Gln	tcg Ser 1115	ttc Phe	aac Asn	gac Asp	tac Tyr	ctc Leu 1120	tgc Cys	gct Ala	gcc Ala	aac Asn	atg Met 1125	ctc Leu	tac Tyr	ccc Pro	20116			
atc Ile	ccc Pro 1130	gcc Ala	aac Asn	gcc Ala	acc Thr	agc Ser 1135	gtg Val	ccc Pro	atc Ile	tcc Ser	att Ile 1140	ccc Pro	tcg Ser	cgg Arg	20161			
aac Asn	tgg Trp 1145	gcc Ala	gcc Ala	ttc Phe	cgg Arg	ggc Gly 1150	tgg Trp	agc Ser	ttc Phe	acc Thr	cgg Arg 1155	ctc Leu	aag Lys	acc Thr	20206			
aag Lys	gag Glu 1160	acc Thr	ccc Pro	tct Ser	ctg Leu	ggc Gly 1165	tcc Ser	ggc Gly	ttc Phe	gat Asp	ccc Pro 1170	tac Tyr	ttc Phe	acc Thr	20251			
tac Tyr	tcg Ser 1175	ggc Gly	tcc Ser	atc Ile	ccc Pro	tac Tyr 1180	ctg Leu	gac Asp	ggc Gly	acc Thr	ttc Phe 1185	tac Tyr	ctc Leu	aac Asn	20296			
cac His	act Thr 1190	ttc Phe	aag Lys	aag Lys	gtc Val	tcc Ser 1195	atc Ile	atg Met	ttc Phe	gac Asp	tcc Ser 1200	tcc Ser	gtc Val	agc Ser	20341			
tgg Trp	ccc Pro 1205	ggc Gly	aac Asn	gac Asp	cgc Arg	ctg Leu 1210	ctg Leu	acc Thr	ccc Pro	aac Asn	gag Glu 1215	ttc Phe	gag Glu	atc Ile	20386			
aag Lys	cgc Arg 1220	acc Thr	gtg Val	gac Asp	ggg Gly	gaa Glu 1225	ggg Gly	tac Tyr	aac Asn	gtg Val	gcc Ala 1230	cag Gln	tgc Cys	aac Asn	20431			
atg Met	acc Thr 1235	aag Lys	gac Asp	tgg Trp	ttc Phe	ctc Leu 1240	atc Ile	cag Gln	atg Met	ctc Leu	agc Ser 1245	cac His	tac Tyr	aac Asn	20476			
atc Ile	ggc Gly 1250	tac Tyr	cag Gln	ggc Gly	ttc Phe	tac Tyr 1255	gtg Val	ccc Pro	gag Glu	ggc Gly	tac Tyr 1260	aag Lys	gac Asp	agg Arg	20521			
atg Met	tac Tyr 1265	tct Ser	ttc Phe	ttc Phe	cgc Arg	aac Asn 1270	ttc Phe	caa Gln	ccc Pro	atg Met	agc Ser 1275	cgc Arg	cag Gln	gtg Val	20566			

gtc gac acc acc acc tac acc gac tac aaa aac gtc acc ctc ccc	20611
Val Asp Thr Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu Pro	
1280 1285 1290	
ttc cag cac aac aac tcg ggg ttc gtg gga tac atg ggc ccc acc	20656
Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr	
1295 1300 1305	
atg cgc gag ggg cag gcc tac ccc gcc aac tac ccc tac ccc ctg	20701
Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu	
1310 1315 1320	
atc ggc aag acc gcc gtg ccc agc ctc acg cag aaa aag ttc ctc	20746
Ile Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu	
1325 1330 1335	
tgc gac cgc acc atg tgg cgc atc ccc ttc tcc agt aac ttc atg	20791
Cys Asp Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met	
1340 1345 1350	
tcc atg ggg gcg ctc acc gac ctg ggg cag aac atg ctg tac gcc	20836
Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala	
1355 1360 1365	
aac tcc gcc cac gcc ctc gac atg acc ttc gag gtg gac ccc atg	20881
Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met	
1370 1375 1380	
gat gag ccc acg ctt ctc tat gtt ctg ttc gaa gtg ttc gac gtc	20926
Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val Phe Asp Val	
1385 1390 1395	
gtg cgc atc cac cag ccg cac cgc ggc gtc atc gag gcc gtc tac	20971
Val Arg Ile His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr	
1400 1405 1410	
ctg cgc acg ccg ttc tcg gcc ggt aac gcc acc acc taa ggagggggcc	21020
Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr	
1415 1420 1425	
gccgacggat gggctccagc gagccggagc tggtcgccat cgcgcgcgac ctgggctgcg	21080
ggccctactt cctgggcacc tttgacaaac gcttcccggg cttcgtggcg ccgcacaagc	21140
tggcctgcgc catcgtcaac accgccggac gcgagaccgg cggcgtccac tggctggccc	21200
tggcctggaa ccccccgcagc cgaacctgct acctcttcga ccccttcggc ttctcggacg	21260
acaggctcag gcagatctac cagttcagat acgaaggcct gctccggcgc agcgccttcg	21320
cctccacccc cgaccactgc gtcaccctcg tcaagtccac ccagaccgtc cagggggcccc	21380
gctcggccgc ctgcggcctc ttctgctgca tggtcctgca cgccttcgtg cgctggccccg	21440
cctcccccctat ggacggcaac cccaccatgg acctccttac gggcgttccc aacagcatgc	21500
ttcagagtcc ccaggtcgag cccaccctcc accgcaacca ggaggaactc tacgccttcc	21560
tggctcggca ctccccctac tttcgccgcc accgcgagcg catagaaaag gccaccgcgt	21620
ttgacaaaat gaacgactag attttctgtg aaaaacactc aataaagcct ttattggttc	21680
accacacgtg cacgcatgca gactttttat ttaaaagggc tccgcctcct cgtcggcgtg	21740
gctggtgggg agggagacgt tgcgatactg caggcgggag ctccatctga actcgggaat	21800
cagcagcttg ggcagggggc cctcgacgtt ctcgctccac agcttgcgca ccagctgcag	21860

ggcgcccagc aggtcgggag cggagatctt gaagtcgcag ttggggccct gggtgccgag 21920
ggagttgcgg tacaccgggt tggcgactg gaacaccagc acgctggggt gctcgatgct 21980
ggccagcgcc gtcttgctgg tcacctcgtc gccgcgcagg gactccgcgt tgctcagcgc 22040
gaaggcggtc agcttgacac gctgccgacc cagcacgggc accccgctcg gctgggtcag 22100
gcagtcgcag cgcataagca tcagcagccg cttctgcccg tgctgcatct tcggatagtc 22160
ggctcgcagc aaggcctcca tctgccggaa ggccgtctgc gccttgctgc cctccgagaa 22220
gaacagcccc caggacttgc cggagaacac gttgttgccg cagctcacgt cttccacgca 22280
gcagcgcgcg tcgtcgttct tcagctgcac cacgctgcgg cccagcgggt tctgcaccac 22340
cttggtcttg ccgggatgtt ctttcagggc ccgctggccg ttctcgctgg tcacgtccat 22400
ctccaccacc tgctccttct ggatcatctc cagcccgtgg tagcagcgca gcacgccctc 22460
ctgctcgggt caccgctgca gccagacggc gcagccgggt ggctccagct gttgaggttt 22520
caccggcg taggtctcca cgtacgcccg caggaagcgg cccatcatct ccacaaaggt 22580
cttctgaccg gtgaaggtca gctgcagccc gcgatgctcc tcgttgagcc acgtctgaca 22640
gatcttgccg tacaccttgc cctgctcggg cagaaacttg aaagcggcct tctcctcggg 22700
ctccacgtgg tacttctcca tcagcgccga catcagctcc atgcccttct cccaggccga 22760
caccagcggc tccgcgcggg gggtcaccac cgccatgcct cgggaagtgc cggggcgctc 22820
atcttctctc tcctcctcgt cttcttcttg aggcggcggt ggccgaggtt gtctcacgaa 22880
tctcttgccg ttggccttct ggacgatctc cacgcccggg tgggtgaacc cgtgggccac 22940
caccacttgc tcctcttctt cttcgtctgc gggcacgact tcgggagagg gaggcggcgg 23000
aggaaccggg gcggccactg cggccatcgc ggcgttcttg cgcgccttct tggggggcag 23060
aggcggcgct tcgcgctccg ggctgggtct ttgcaggtag ggcgtgatgg tgtgggaggt 23120
ggggcgctct ggctgacggc cggccatgct gatgcttgac tcctaggcga aaagatggag 23180
gaggatctta gacagccgca gcccgtctcc gaaaccttaa cccccccg ctctgaggtc 23240
ggcgccggcg agctagacat gcaacgggag gaggaggagg acgtgcgagt ggagcaagac 23300
ccgggctacg tgacgccgca cgaggacggc gaggagccgc aggcaccggc gccaacgctc 23360
agcgaagccg actacctggg aggggaggag gacgtgctgc tgaagcacct ggcgcggcag 23420
agcaccatcg tgcaggaggc cctcaaggag cgcgaggagg tcccgtgac ggtggaggag 23480
ctcagccggg cctacgaagc caacctcttc tcgcccggg tgcccccaa gaagcaggcc 23540
aacggcacct gcgagcccaa cccccgcctc aacttctacc ccgtctttgc ggtgcccag 23600
gcgctggcca cctatcacat cttcttcaag aaccagcgca tccccctctc gtgccgcgcc 23660
aaccgcacc gcgccgacc cctcctgcat ctccgagccg gcgccgcat acctgagatc 23720
gcctccctgg aggaagtccc caagatcttc gaaggtctcg gcaaggacga gaagcgcgcg 23780
gcaaacgctc tggaaaagaa cgagagcgag ggtcagaacg tgctgggtcga gctggaaggc 23840
gacaacgcgc gtctggccgt gctcaaacgc accatcgaag tctccactt cgcctacccc 23900
gcgctcaacc ttcccccaa ggtcatgcgc tcggtcatgg atcagctgct catcaagcgc 23960

gccgagcccc tcgagaacga ctccgaggtg gattccgagg acggaacc cgtggtctcg 24020
gacgaggagc tcgcgcgctg gctgggcacg caggaccccc ccgagttgca agagcggcgc 24080
aagatgatga tggcggccgt gctggtcacc gccgagctcg agtgcctgca gcgcttcttc 24140
gcccagcccc agaccctgcg caaggctcag gagtccctgc actacgcctt ccgccacggc 24200
tacgtgcgcc aggccctgcaa gatctccaac gtggagctta gcaacctggc ctctacatg 24260
ggcatcctgc acgagaaccg cctcgggcag aacgtcctcc actgcaccct gaccggggag 24320
gcccgcgcgc actacgtccg cgactgcac tacctctttc tcaccctcac ctggcagacc 24380
gccatggggg tctggcagca gtgtctggag gagcgcaacc ttcgcgagct cgacaagcta 24440
ctgagccgcg agcgcgcga gctctggacg gctttcagcg agcgcaccgc cgcctgccgt 24500
ctggccgacc tcatcttccc cgagcgactc aggcaaacc tccagaacgg cctgcccgac 24560
ttgtcagcc agagcatgct gcaaaacttt cgctccttca tcctggagcg atccggcatc 24620
ttgcccgcga tgagctgcgc cctgccctcc gatttcgtcc ccctctatta tcgcgagtgc 24680
ccccgcgcgc tctggagcca ctgctacctg ctgctctgg ccaactacct cgcaccac 24740
tccgacctca tggaagactc cagcggcgag gggctgctgg agtgccactg ccgctgcaac 24800
ctctgcacc cccaccgctc gctggtctgc aacaccgagc tgctcagcga gacgcaagtg 24860
atcggctacct ttgagatcca gggaccagag gggccggagg gtgcttccaa cctcaagctc 24920
agcccggcgc tctggacttc cgcctacctg cgcaaattta tccccgagga ctatcacgcc 24980
caccagatcc aattctacga agaccaatcg cgacccccca aagccccct caccgctgt 25040
gtcatcacc agagccagat tctggcccaa ttgcaagcca tccagcaggc ccgccaagag 25100
ttctctctga aaaagggtca cggggtctat ctggaccccc agaccggcga ggaactcaac 25160
accccgtcac cctccgcgc cgcttcgtgc cgccgcaga accatgccgc ccaaagggaa 25220
caagcaggcc atcgcccagc ggcgggcca gaagcagcaa gagctccagg agcagtggga 25280
cgaggagtcc tgggacagcc aggcggagga agtctcagac gaggaggagg acatggagag 25340
ctgggacagc ctagacgagg aggaggaggc cgaggagcta gaggacgagc ctctcagga 25400
ggaagagccc agcagcgccg cggcaccatc ggcttccaaa gaagcggctc ggagccggcc 25460
ggccccgaag cagcagaagc agcaacagcc gccaccgtcg ccccgacgc caccaccagg 25520
ctcactcaaa gccagccgta ggtgggacgc ggtgtccatc gcgggatcgc ccaaagcccc 25580
agtcggtaag ccacccgggc ggtcgcggcg ggggtactgt tcctggcgcc ccacaagag 25640
caagatcgtc gcctgcctcc agcactgccg gggcaacatc tccttcgcgc ggcgctactt 25700
gtcttccac gacggggtgg cgggtccgcg caacgtctc tactattacc gtcattctta 25760
cagcccctac gagacagaag gcccggcctc cgcgtaagac cagccgccag acggtctcct 25820
ccgccatcgc gaccgcccag gactcggccg ccacgcagga gctcagaaaa cgcatctttc 25880
ccaccctgta tgctatcttc cagcagagcc gcggccagca gctggaactg aaagtaaaaa 25940
accgctccct gcgttcgctc acccgagct gtctgtacca caggaggga gaccaactgc 26000

agcgcacgct cgaggacgcc gaggcactgt tcaataaata ctgctcggtg tctcttaagg 26060
 actgaaagcc cgcgcttttt cagaggctca ttacgtcatc atcatcatga gcaaggacat 26120
 tcccacgcct tacatgtgga gctaccagcc gcagatggga ctggcgggccg gcgcctccca 26180
 ggattactcc agtcgcatga actggctgag tgccggcccc cacatgatcg ggcgggtcaa 26240
 tgggattcgt gccacccgca atcagatact gctggaacag gccgccctca cctccacccc 26300
 gcgacgtcag ctgaacccgc ccgcttggcc cgccgcccag gtgtaccagg aaaaccccg 26360
 cccgaccaca gtcctcctgc cacgcgacgc ggaggccgaa gtccagatga ctaactccgg 26420
 ggcgcaatta gcgggcgggc cccgccacgt cgtcgctccc gggtagagag gtcggcccg 26480
 accctacccc tccggcccta taaagaggct gatcattcga ggccgaggta tccagctcaa 26540
 cgacgaggtg gtgagctcct cgaccggtct tcggcccgac ggagtcttcc agcttgagg 26600
 cgccggccgc tcttccttca cactcgcca ggcctacctg acgctccaga gctcttcctc 26660
 ccagcctcgc tccggcggca tcggcaccct ccagttcgtg gaggagtctg tgccctcgg 26720
 ctacttcaac ccgttctccg gctctcccg ccgctacccg gacagcttca tccccaaacta 26780
 cgacgcggtg agcgaatccg tggacggcta cgattgatga ccgatggtgc ggccgtaact 26840
 gcgcggcggc aacatctgca tcaactgcat cgtcctcggg gcttcgcccg ggaggcctgt 26900
 gagttcatct acttcagct cgccccggac cagcttcagg gcccttcgca cggcgtaag 26960
 ctcgtgatag aggaagagct cgagagtagc tgcctgcgct gttttacctc gcgccccatc 27020
 ctagtgcaga gggaacgcgg taggaccacc ctcaccctct actgcatctg tgactccccg 27080
 gaattacatg aagatctgtg ttgccttcta tgtgccgaac aataaccct cttgtaacta 27140
 cctacatcca caataaacca gaatttgga actcctttcg tttgtttgca g atg aaa 27197
 Met Lys

cgc gcc cgc ctc gac gac gac ttc aac ccc gtc tac ccc tat gac 27242
 Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr Asp
 1430 1435 1440

act ccc aac gct ccc tct gtt ccc ttc atc act cct ccc ttc gtc 27287
 Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val
 1445 1450 1455

tcc tcg gac ggc ttg caa gaa aaa cca ccc gga atg ctc agt ctc 27332
 Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu
 1460 1465 1470

aac tac caa gat cct att acc acc caa aac ggg gca tta act cta 27377
 Asn Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu
 1475 1480 1485

aag ctt ggc agc gga ctg aac ata aac caa gat ggg gaa ctt acc 27422
 Lys Leu Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr
 1490 1495 1500

tca gac gcc agc gtt ctc gtc act ccc ccc att aca aaa gcc aac 27467
 Ser Asp Ala Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn
 1505 1510 1515

aac aca ata ggc cta gcc ttc aat gca cct ctt acc ttg caa agc 27512
 Asn Thr Ile Gly Leu Ala Phe Asn Ala Pro Leu Thr Leu Gln Ser
 1520 1525 1530

gat act tta aat ctt gct tgt aac gcc cca ctt acc gtg caa gac	27557
Asp Thr Leu Asn Leu Ala Cys Asn Ala Pro Leu Thr Val Gln Asp	
1535 1540 1545	
aat agg ttg gga ata aca tac aac tct ccc ctc acc ttg caa aac	27602
Asn Arg Leu Gly Ile Thr Tyr Asn Ser Pro Leu Thr Leu Gln Asn	
1550 1555 1560	
agc gaa ctt gcc cta gcg gtc acc ccg cct ctt gac act gcc aat	27647
Ser Glu Leu Ala Leu Ala Val Thr Pro Pro Leu Asp Thr Ala Asn	
1565 1570 1575	
aac aca ctt gcg ctt aaa acc gcc cgg cct ata att aca aac tct	27692
Asn Thr Leu Ala Leu Lys Thr Ala Arg Pro Ile Ile Thr Asn Ser	
1580 1585 1590	
aat aac gag ctt aca ctc tcc gct gat gct ccc cta aac acc agc	27737
Asn Asn Glu Leu Thr Leu Ser Ala Asp Ala Pro Leu Asn Thr Ser	
1595 1600 1605	
acg ggt acc ctc cgc cta caa agc gca gca cca ctg ggg cta gtt	27782
Thr Gly Thr Leu Arg Leu Gln Ser Ala Ala Pro Leu Gly Leu Val	
1610 1615 1620	
gac caa acc ctg cga gtg ctt ttt tct aac cca ctc tac ttg caa	27827
Asp Gln Thr Leu Arg Val Leu Phe Ser Asn Pro Leu Tyr Leu Gln	
1625 1630 1635	
aac aac ttt ctc tca cta gcc att gaa cgc cca ttg gct tta act	27872
Asn Asn Phe Leu Ser Leu Ala Ile Glu Arg Pro Leu Ala Leu Thr	
1640 1645 1650	
acc act ggt tct atg gct atg cag att tcc caa cca tta aaa gtg	27917
Thr Thr Gly Ser Met Ala Met Gln Ile Ser Gln Pro Leu Lys Val	
1655 1660 1665	
gaa gac gga agc tta agc ttg agc att gaa agc cct cta aat cta	27962
Glu Asp Gly Ser Leu Ser Leu Ser Ile Glu Ser Pro Leu Asn Leu	
1670 1675 1680	
aaa aac gga aat ctt act tta gga acc caa agt ccc cta act gtc	28007
Lys Asn Gly Asn Leu Thr Leu Gly Thr Gln Ser Pro Leu Thr Val	
1685 1690 1695	
act ggt aac aac ctc agc ctt aca aca aca gcc cca tta acg gtt	28052
Thr Gly Asn Asn Leu Ser Leu Thr Thr Thr Ala Pro Leu Thr Val	
1700 1705 1710	
cag aac aac gct cta gcc ctc tca gtg tta ctg ccg ctt aga cta	28097
Gln Asn Asn Ala Leu Ala Leu Ser Val Leu Leu Pro Leu Arg Leu	
1715 1720 1725	
ttt aat aac acc tca ctg gga gtg gca ttc aac cca ccc att tct	28142
Phe Asn Asn Thr Ser Leu Gly Val Ala Phe Asn Pro Pro Ile Ser	
1730 1735 1740	
tca gca aac aac ggg ctg tct ctt gac att gga aat ggc ctt aca	28187
Ser Ala Asn Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr	
1745 1750 1755	
ctg caa tac aac agg ctc gta gtg aac att ggc ggc ggg cta cag	28232
Leu Gln Tyr Asn Arg Leu Val Val Asn Ile Gly Gly Gly Leu Gln	
1760 1765 1770	
ttt aac aac ggt gct att acc gct tcc ata aat gca gct ctg ccg	28277
Phe Asn Asn Gly Ala Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro	
1775 1780 1785	

ttg	cag	tat	tcc	aat	aac	cag	ctt	tct	ctt	aat	att	gga	ggc	ggg	28322
Leu	Gln	Tyr	Ser	Asn	Asn	Gln	Leu	Ser	Leu	Asn	Ile	Gly	Gly	Gly	
		1790					1795					1800			
ctg	cga	tac	aac	ggc	act	tac	aaa	aat	tta	gcc	gtc	aaa	acc	gac	28367
Leu	Arg	Tyr	Asn	Gly	Thr	Tyr	Lys	Asn	Leu	Ala	Val	Lys	Thr	Asp	
		1805					1810					1815			
tct	ttt	agg	ggg	ctt	gaa	att	gac	agt	aat	cag	ttc	ctg	gtg	cca	28412
Ser	Phe	Arg	Gly	Leu	Glu	Ile	Asp	Ser	Asn	Gln	Phe	Leu	Val	Pro	
		1820					1825					1830			
aga	ctg	ggg	tct	ggg	cta	aag	ttt	gat	caa	tat	ggg	tac	att	agc	28457
Arg	Leu	Gly	Ser	Gly	Leu	Lys	Phe	Asp	Gln	Tyr	Gly	Tyr	Ile	Ser	
		1835					1840					1845			
gtc	ata	cct	cca	act	ggt	acg	cca	aca	aca	ctt	tgg	act	aca	gca	28502
Val	Ile	Pro	Pro	Thr	Val	Thr	Pro	Thr	Thr	Leu	Trp	Thr	Thr	Ala	
		1850					1855					1860			
gac	cct	tct	ccc	aac	gct	act	ttt	tac	gac	agc	tta	gat	gct	aag	28547
Asp	Pro	Ser	Pro	Asn	Ala	Thr	Phe	Tyr	Asp	Ser	Leu	Asp	Ala	Lys	
		1865					1870					1875			
gta	tgg	ctg	gcc	tta	gta	aaa	tgc	aac	ggc	atg	ggt	aat	gga	acc	28592
Val	Trp	Leu	Ala	Leu	Val	Lys	Cys	Asn	Gly	Met	Val	Asn	Gly	Thr	
		1880					1885					1890			
ata	gcc	ata	aag	gct	tta	aaa	ggt	act	ctg	ctc	caa	cct	acg	gct	28637
Ile	Ala	Ile	Lys	Ala	Leu	Lys	Gly	Thr	Leu	Leu	Gln	Pro	Thr	Ala	
		1895					1900					1905			
agt	ttt	att	tct	ttt	ggt	atg	tat	ttt	tac	agc	aat	ggc	acc	aga	28682
Ser	Phe	Ile	Ser	Phe	Val	Met	Tyr	Phe	Tyr	Ser	Asn	Gly	Thr	Arg	
		1910					1915					1920			
aga	act	aac	tac	ccc	acg	ttt	gaa	aat	gaa	ggc	ata	cta	gct	agt	28727
Arg	Thr	Asn	Tyr	Pro	Thr	Phe	Glu	Asn	Glu	Gly	Ile	Leu	Ala	Ser	
		1925					1930					1935			
agt	gct	aca	tgg	ggg	tat	cgt	caa	gga	aac	tcg	gca	aac	acc	aac	28772
Ser	Ala	Thr	Trp	Gly	Tyr	Arg	Gln	Gly	Asn	Ser	Ala	Asn	Thr	Asn	
		1940					1945					1950			
gtc	acc	agt	gcc	ggt	gaa	ttt	atg	cct	agc	tcc	aca	aga	tat	cct	28817
Val	Thr	Ser	Ala	Val	Glu	Phe	Met	Pro	Ser	Ser	Thr	Arg	Tyr	Pro	
		1955					1960					1965			
gtt	aac	aag	ggg	act	gag	ggt	cag	aac	atg	gaa	ctc	acc	tac	act	28862
Val	Asn	Lys	Gly	Thr	Glu	Val	Gln	Asn	Met	Glu	Leu	Thr	Tyr	Thr	
		1970					1975					1980			
ttc	ttg	cag	gga	gac	ccc	act	atg	gcc	ata	tca	ttt	caa	gct	att	28907
Phe	Leu	Gln	Gly	Asp	Pro	Thr	Met	Ala	Ile	Ser	Phe	Gln	Ala	Ile	
		1985					1990					1995			
tat	aac	cat	gct	ttg	gaa	ggg	tac	tct	tta	aaa	ttt	acc	tgg	cga	28952
Tyr	Asn	His	Ala	Leu	Glu	Gly	Tyr	Ser	Leu	Lys	Phe	Thr	Trp	Arg	
		2000					2005					2010			
gtt	cgc	aac	agg	gaa	cgc	ttt	gat	atc	ccc	tgc	tgt	tct	ttt	tct	28997
Val	Arg	Asn	Arg	Glu	Arg	Phe	Asp	Ile	Pro	Cys	Cys	Ser	Phe	Ser	
		2015					2020					2025			
tac	ata	acg	gaa	gaa	taa	acactg	tttt	tctttt	caat	gttttt	tattc				29045
Tyr	Ile	Thr	Glu	Glu											
		2030													
tgctttttta	cacagttcga	accgtcagac	tccctcccc	cttccacttc	acccggtaca										29105

cctcccgtc cccctggatc gctgcgtaca actgcagttt ggtgttcaga cacgggttct 29165
 taggtgacag tatccacacg gcctctttgc cggccaggcg ctggtccgta atgctcacia 29225
 atccctccga cacgtcctcc agacacacgg tggaatccaa ggcgcccgtc tacaaaacaa 29285
 acacagtcac gctctccacg ggttctctcc tcggtcgtac tgcgccagcg tgaacgggag 29345
 atggtgctcc atcagggtc gcagcaaccg ctgtcggcgc ggctcaccca ggctccggcg 29405
 aaaagcggc cgtctgggag tgctattcaa aaaacgcacc gcctttatca acagtctcct 29465
 cgtgcggcgg gcgcagcagc gcacctggat ctctgtcagg tctttacaat aggtacagcc 29525
 catcaccacc atgttgttta aaatcccaaa gctaaacacg ctccacccaa atgacatgaa 29585
 ttccagcacc gccgcggcgt ggccatcata caatatgcgg aggtaaatca ggtgccgccc 29645
 cctaatacaa acgtcccca tatacatcac ctcttaggc agttgataat taaccacctc 29705
 ccggtaccag ggaaacctca cgtttactaa agcccaaac accaacattt taaaccagtt 29765
 agccagcacc acccctccc ccttacctg cagcgacccc ggctgtttac aatgacagt 29825
 aatcacccac ctctcatacc ccctaatac ctggcggtggc tccacatcta tagtagcaca 29885
 gcacacgcac accctcatgt aatgcttcat cacaaatctt tccaagggg ttagtatcat 29945
 gtcccagggt acggggcact cctgcagcac ggtgaaagg acgcaggcg gaacagtcct 30005
 cacctcggac acataatgca tattcagatg ttcacactct aaaaccccg ggcttccctc 30065
 caacgcagcc actggcaagt tctcagaggg tggtgtaagg cgggtggtgct gatagggact 30125
 caatctgtgt cgacaccgtc tgtcgcgttg catcgtagac caacgcttg cgacaccgct 30185
 cgtacttcgc ccaaagaaaa cgggtgcgac gccaacacac ttccgcgtac cgtgggttcc 30245
 gactcagac tcgctcagtt ctcaacgcac aatgcagcca ttctgtaat ccacacaaca 30305
 gtcgctcggc ttcaaagag atgtgcacct cgtatcttat aacgtcccga tatatatcca 30365
 agcaggcagt caggggcact tgcaaccagt gcacgcaggc ggactgatcg cgacacactg 30425
 gaggtggagg gagagacgga agaggcatgt tactccagac ggtcgaaaag cggatcaaag 30485
 tgcagatcgc gaagatggca gcgatccccg ccgctacgct ggtgatagat cacagccagg 30545
 tcaaacataa tgcggttttc caaatgacct attaccgcct ccaccagagc cgccacgcgc 30605
 acttccagaa acaccagcac ggctacggca ttctcctcaa aatcttcaa cattaagctg 30665
 catgattgaa tcaccccaa ataattctcc tccttccatt ctgcgaaaat ttgagtaaaa 30725
 acctctcgca gattagctcc gtggcggttca aaaagggtcac ttagagcgcc ctccaccgcc 30785
 atgcgcaagc acaccctcat gattgaaaaa tgccagtctc ctgaaccacc tgcagttgat 30845
 ttaaaagacc tatattagga tcaattccac tctcccgcag ctccacgcgt agcattagct 30905
 gcaaaaagtc atttaaatct tcgcaaaacta gcgcggtaag ctgcgcccg ggaattaggt 30965
 ctgaagcagt caccacacac ataatttcca gtgaaggagt cagtctaagc agcaaaaagc 31025
 cgcagtagca gtgttgaaaa ggaggggtca cgcaatgtaa catatgcagc caaaaatctc 31085
 caagggtgtc gtgcataaac tccaccactg aaaagtccaa atcatgtaaa tatgccatca 31145

ccgcctcagg aaccaccacg gacacaaaaa cgggccgtag caaatacatg gtgtcctgca 31205
 aagcaaaaac acattttatac catagaggcg cgaattactt ggggaaaaat cactcgctcc 31265
 aaaactaaac aggccaccgt ctgaccgcgc cagccataaa aaaagcgggtt cgaatgatta 31325
 aaaagaataa tagacacctc ccaccaggta ctcggctgca actcgtgcgc ccctatcaaa 31385
 accccgcgga cgttcatgtc ggccatagaa aaaatgcggc ccaaatatcc caccggaatc 31445
 tccacggcca gctgcagtga tagcaaaaga acgcatgag gagcaatcac aaaattttca 31505
 ggcgataaaa gcacataaag gttagaatag ccctgctgca caggtaataa agcccgcgag 31565
 ctcagcaa at gcacataaac cgcttcagcc atcccgtctt accgcgaaca aaaggctcac 31625
 agtacacagt tactcaaccc acacgccaca cagtatttat acactcctca atcgccacgt 31685
 caccgcggcc gaacaaactc caaaagtcca aaaagtccaa aacgcccgcg taaaagcccg 31745
 ccaaaacagc acttcctcat ttactctccc acagtacgtc acttccgccg cgcccgcgcg 31805
 cctcgccccg ccctcaccct cgcgctccac cccgcgcccc acgtcagact cccacccccg 31865
 cccgcgcccc cgctcatccg accccaccct cactccaccc ctaacccccg ctctcatta 31925
 tcatattggc accgttccca aataaggat attatgatga tg 31967

<210> 13
 <211> 508
 <212> PRT
 <213> Simian adenovirus

<400> 13

Met Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly
1 5 10 15

Pro Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala
20 25 30

Thr Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu
35 40 45

Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr
50 55 60

Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Ile Ala Ser Leu
65 70 75 80

Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn
85 90 95

Asn Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp
100 105 110

Glu Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn
115 120 125

Met Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg
Page 124

130

135

140

Leu Met Val Glu Lys Val Asn Lys Glu Thr Asn Ala Pro Arg Tyr Glu
 145 150 155 160

Trp Phe Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser Glu Thr Met Thr
 165 170 175

Ile Asp Leu Met Asn Asn Ala Ile Val Asp Asn Tyr Leu Glu Val Gly
 180 185 190

Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr
 195 200 205

Arg Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro
 210 215 220

Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro
 225 230 235 240

Gly Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly
 245 250 255

Ile Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu
 260 265 270

Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr
 275 280 285

Glu Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Gln Glu Ile Arg Gly
 290 295 300

Asp Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu
 305 310 315 320

Lys Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr
 325 330 335

Asn Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro
 340 345 350

Glu Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Thr Asp Val Thr
 355 360 365

Cys Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp
 370 375 380

Pro Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val
 385 390 395 400

Gly Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln
 405 410 415

Ala Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val
420 425 430

Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro
435 440 445

Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly
450 455 460

Thr Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile
465 470 475 480

Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly
485 490 495

Ile Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe
500 505

<210> 14
<211> 917
<212> PRT
<213> Simian adenovirus

<400> 14

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala
1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
20 25 30

Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
50 55 60

Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65 70 75 80

Lys Val Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
115 120 125

Ala Pro Asn Pro Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val
130 135 140

Arg Gly Gln Ala Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile
 145 150 155 160
 Gln Val Ala Thr Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu
 165 170 175
 Tyr Gln Pro Glu Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala
 180 185 190
 Ser Asp Ser Asp Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro
 195 200 205
 Met Phe Pro Cys Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly
 210 215 220
 Gly Gln Gly Thr Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ser
 225 230 235 240
 Ala Thr Ser Thr Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile
 245 250 255
 Glu Ala Pro Asp Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr
 260 265 270
 Thr Thr Ser Ser Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg
 275 280 285
 Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr
 290 295 300
 Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu
 305 310 315 320
 Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln
 325 330 335
 Leu Met Leu Asp Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp
 340 345 350
 Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn
 355 360 365
 His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly
 370 375 380
 Ser Leu Val Thr Glu Thr Tyr Thr Gly Leu Ser Pro Gln Asn Gly Ser
 385 390 395 400
 Asn Thr Trp Thr Thr Asp Ser Thr Thr Tyr Ala Thr Arg Gly Val Glu
 405 410 415
 Ile Gly Ser Gly Asn Met Phe Ala Met Glu Ile Asn Leu Ala Ala Asn
 Page 127

420

425

430

Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp
 435 440 445

Glu Tyr Lys Leu Thr Pro Asp Asn Ile Thr Leu Pro Asp Asn Lys Asn
 450 455 460

Thr Tyr Asp Tyr Met Asn Gly Arg Val Ala Ala Pro Ser Ser Leu Asp
 465 470 475 480

Thr Tyr Val Asn Ile Gly Ala Arg Trp Ser Pro Asp Pro Met Asp Asn
 485 490 495

Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser
 500 505 510

Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro
 515 520 525

Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr
 530 535 540

Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser
 545 550 555 560

Ser Leu Gly Asn Asp Leu Arg Val Asp Gly Ala Ser Val Arg Phe Asp
 565 570 575

Ser Ile Asn Leu Tyr Ala Asn Phe Phe Pro Met Ala His Asn Thr Ala
 580 585 590

Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe
 595 600 605

Asn Asp Tyr Leu Cys Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn
 610 615 620

Ala Thr Ser Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe
 625 630 635 640

Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu
 645 650 655

Gly Ser Gly Phe Asp Pro Tyr Phe Thr Tyr Ser Gly Ser Ile Pro Tyr
 660 665 670

Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile
 675 680 685

Met Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr
 690 695 700

Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn
705 710 715 720

Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Ile Gln Met Leu
725 730 735

Ser His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr
740 745 750

Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg
755 760 765

Gln Val Val Asp Thr Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu
770 775 780

Pro Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr
785 790 795 800

Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile
805 810 815

Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu Cys Asp
820 825 830

Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly
835 840 845

Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His
850 855 860

Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu
865 870 875 880

Leu Tyr Val Leu Phe Glu Val Phe Asp Val Val Arg Ile His Gln Pro
885 890 895

His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala
900 905 910

Gly Asn Ala Thr Thr
915

<210> 15
<211> 607
<212> PRT
<213> Simian adenovirus

<400> 15

Met Lys Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr
1 5 10 15

Asp Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val
 20 25 30
 Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu Asn
 35 40 45
 Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu Lys Leu
 50 55 60
 Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr Ser Asp Ala
 65 70 75 80
 Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn Asn Thr Ile Gly
 85 90 95
 Leu Ala Phe Asn Ala Pro Leu Thr Leu Gln Ser Asp Thr Leu Asn Leu
 100 105 110
 Ala Cys Asn Ala Pro Leu Thr Val Gln Asp Asn Arg Leu Gly Ile Thr
 115 120 125
 Tyr Asn Ser Pro Leu Thr Leu Gln Asn Ser Glu Leu Ala Leu Ala Val
 130 135 140
 Thr Pro Pro Leu Asp Thr Ala Asn Asn Thr Leu Ala Leu Lys Thr Ala
 145 150 155 160
 Arg Pro Ile Ile Thr Asn Ser Asn Asn Glu Leu Thr Leu Ser Ala Asp
 165 170 175
 Ala Pro Leu Asn Thr Ser Thr Gly Thr Leu Arg Leu Gln Ser Ala Ala
 180 185 190
 Pro Leu Gly Leu Val Asp Gln Thr Leu Arg Val Leu Phe Ser Asn Pro
 195 200 205
 Leu Tyr Leu Gln Asn Asn Phe Leu Ser Leu Ala Ile Glu Arg Pro Leu
 210 215 220
 Ala Leu Thr Thr Thr Gly Ser Met Ala Met Gln Ile Ser Gln Pro Leu
 225 230 235 240
 Lys Val Glu Asp Gly Ser Leu Ser Leu Ser Ile Glu Ser Pro Leu Asn
 245 250 255
 Leu Lys Asn Gly Asn Leu Thr Leu Gly Thr Gln Ser Pro Leu Thr Val
 260 265 270
 Thr Gly Asn Asn Leu Ser Leu Thr Thr Thr Ala Pro Leu Thr Val Gln
 275 280 285
 Asn Asn Ala Leu Ala Leu Ser Val Leu Leu Pro Leu Arg Leu Phe Asn

290

295

300

Asn Thr Ser Leu Gly Val Ala Phe Asn Pro Pro Ile Ser Ser Ala Asn
 305 310 315 320

Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr Leu Gln Tyr Asn
 325 330 335

Arg Leu Val Val Asn Ile Gly Gly Gly Leu Gln Phe Asn Asn Gly Ala
 340 345 350

Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro Leu Gln Tyr Ser Asn Asn
 355 360 365

Gln Leu Ser Leu Asn Ile Gly Gly Gly Leu Arg Tyr Asn Gly Thr Tyr
 370 375 380

Lys Asn Leu Ala Val Lys Thr Asp Ser Phe Arg Gly Leu Glu Ile Asp
 385 390 395 400

Ser Asn Gln Phe Leu Val Pro Arg Leu Gly Ser Gly Leu Lys Phe Asp
 405 410 415

Gln Tyr Gly Tyr Ile Ser Val Ile Pro Pro Thr Val Thr Pro Thr Thr
 420 425 430

Leu Trp Thr Thr Ala Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser
 435 440 445

Leu Asp Ala Lys Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val
 450 455 460

Asn Gly Thr Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro
 465 470 475 480

Thr Ala Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr
 485 490 495

Arg Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser
 500 505 510

Ser Ala Thr Trp Gly Tyr Arg Gln Gly Asn Ser Ala Asn Thr Asn Val
 515 520 525

Thr Ser Ala Val Glu Phe Met Pro Ser Ser Thr Arg Tyr Pro Val Asn
 530 535 540

Lys Gly Thr Glu Val Gln Asn Met Glu Leu Thr Tyr Thr Phe Leu Gln
 545 550 555 560

Gly Asp Pro Thr Met Ala Ile Ser Phe Gln Ala Ile Tyr Asn His Ala
 565 570 575

Leu Glu Gly Tyr Ser Leu Lys Phe Thr Trp Arg Val Arg Asn Arg Glu
580 585 590

Arg Phe Asp Ile Pro Cys Cys Ser Phe Ser Tyr Ile Thr Glu Glu
595 600 605

<210> 16
<211> 24
<212> DNA
<213> Artificial

<220>
<223> synthetic oligomer

<400> 16
gcgacgggcc gacgctgccc ggct

24

<210> 17
<211> 4
<212> PRT
<213> Artificial

<220>
<223> artifical

<400> 17

Arg Arg Ala Ser
1

<210> 18
<211> 24
<212> DNA
<213> Artificial

<220>
<223> synthetic oligomer

<400> 18
gcggcgcgcc gacgctgccc ggct

24

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.